Intertemporal preferences of individual and collective decision-makers

Nina Anchugina

Department of Mathematics
The University of Auckland
New Zealand

May 2017

A thesis submitted in fulfilment of the requirements of Doctor of Philosophy in Mathematics.

Abstract

This thesis contributes to theoretical modelling of time preferences of individual and collective decision-makers. In the context of an individual decision-maker we study two problems: axiomatization of time preferences and the effect of delay on the ranking of sequences of dated outcomes. In the context of a group of decision-makers we investigate the problem of aggregation of time preferences.

First, we provide a new axiomatic foundation for exponential, quasi-hyperbolic and semi-hyperbolic discounting when preferences are expressed over streams of consumption lotteries. The key advantage of our axiomatic system is its simplicity and its use of a common framework for finite and infinite time horizons.

Second, we analyse preferences with the property that the ranking of two sequences of dated outcomes can switch from one strict ranking to the opposite at most once as a function of some common delay – the "one-switch" property of Bell [12]. We demonstrate that time preferences satisfy the one-switch property if and only if the discount function is either the sum of exponentials or linear times exponential. This is a revision of Bell's result [12], who claimed that the only discount functions compatible with the one-switch property are sums of exponentials. We also show that linear times exponential discount functions exhibit increasing impatience in the sense of Takeuchi [77]. To the best of our knowledge, linear times exponential discount functions have not been used in the context of time preference before.

Finally, we study the problem of aggregating time preferences when individual time preferences exhibit decreasing impatience. If decision-makers have the same level of decreasing impatience, our result proves that the aggregate discount function is strictly more decreasingly impatient than each of individual discount functions. This is a generalization of Prelec's and Jackson and Yariv's results on the aggregation of discount functions [46, 63]. We also analyse the situation in which the aggregation problem arises because of some uncertainty about the discount function. In this context we prove the analogue of Weitzman's influential result [81], showing that if a decision-maker is uncertain about her hyperbolic discount rate, then long-term costs and benefits will be discounted at a hyperbolic discount rate which is the probability-weighted harmonic mean of the possible hyperbolic discount rates.

Acknowledgments

First of all, I would like to deeply thank my supervisor Associate Professor Matthew Ryan for his endless patience, constant support, encouragement and dedication. Your deep knowledge of logic drove me to present the arguments in a clear and concise way. Your attention to detail did not let my proofs go completely astray despite my numerous attempts to do so. It has been a pleasure to work with you over the last three and a half years, and I owe you a huge debt of gratitude for guiding me to the completion of this thesis.

I am thankful to my supervisor Professor Steven Galbraith for his support and especially for taking on the administrative side of things.

To my co-author Professor Arkadii Slinko, thank you for collaboration and contribution to presenting and organizing material in Chapter 3, for your thoughtful advice and comments.

Professor Simon Grant, thank you for hosting me at Research School of Economics at Australian National University and for your advice and encouragement to publish my work. With your inspiration the results from Chapter 2 of this thesis have been accepted for publication in Theory and Decision [6] and the results from Chapter 3 have been submitted to a journal.

To Doctor Vladimir Petkov and Doctor Paul Calcott, thank you for your insight and interest in my work and hosting me at School of Economics and Finance at Victoria University of Wellington on two occasions to present and share my work.

My sincere gratitude goes to Vivien Kirk for her kindness and mentoring through the last year of my PhD.

Thank you to my fellow PhD students and friends Andrus, Rebecca, Susana, Iresha and Pascal for sharing the good times and the bad, our discussions, dancing and hiking adventures. I would also like to thank my friend Varvara, whose footprints I followed all the way from Siberia to New Zealand, for her support, sense of humour and certain adventures.

This work would not have been possible without University of Auckland Doctoral

Scholarship. I am grateful to the Department of Mathematics for providing the opportunity to study away from my home country at a world class university, hosting my PhD and providing the travel funds.

Finally I wish to thank my family for their unconditional support. Special thanks go to my dear Malcolm for his love, sharing his life with me and always willing to suffer through multiple practice talks on my PhD research.

Contents

List of Figures

List of Tables

THE UNIVERSITY OF AUCKLAND
SCHOOL OF GRADUATE STUDIES

Co-Authorship Form

Graduate Centre
The ClockTower – East Wing
22 Princes Street, Auckland
Phone: +64 9 373 7599 ext 81321
Fax: +64 9 373 7610
Email: postgraduate@auckland.ac.nz
www.postgrad.auckland.ac.nz

This form is to accompany the submission of any PhD that contains published or unpublished co-authored work. **Please include one copy of this form for each co-authored work.** Completed forms should be included in all copies of your thesis submitted for examination and library deposit (including digital deposit), following your thesis Acknowledgements. Co-authored works may be included in a thesis if the candidate has written all or the majority of the text and had their contribution confirmed by all co-authors as not less than 65%.

Please indicate the chapter/section/pages of this thesis that are extracted from a co-authored work and give the title and publication details or details of submission of the co-authored work.

The content of Section 1.1.5 and Chapter 3 is co-authored work:

Anchugina, N., Ryan, M., & Slinko, A. (2016). Aggregating time preferences with decreasing impatience. arXiv preprint arXiv:1604.01819.

Nature of contribution by PhD candidate	Performed the investigation, the original ideas for the proofs; generated figures; wrote, proofread and edited the text.
Extent of contribution by PhD candidate (%)	75

CO-AUTHORS

Name	Nature of Contribution
Matthew Ryan	Advised on the investigation; some details in the proofs of theorems and propositions; proofread and edited the text.
Arkadii Slinko	Advised on the investigation; the initial idea of Theorem 3.28, some details in the proofs of theorems and propositions; proofread and edited the text.

Certification by Co-Authors

The undersigned hereby certify that:

- the above statement correctly reflects the nature and extent of the PhD candidate's contribution to this work, and the nature of the contribution of each of the co-authors; and
- that the candidate wrote all or the majority of the text.

Name	Signature	Date
Matthew Ryan	M Ryan	18/12/2016
Arkadii Slinko	ASlinko	14/12/2016

Last updated: 19 October 2015

1
Introduction

The focus of this thesis is discounting. Discounting arises in situations where intertemporal choice is involved, such as the necessity to compare alternatives which occur in different periods of time. For example, there is always a time dimension to decisions about investment projects, savings, pensions, mortgage. Such decisions may be made at the individual level or by a group of individuals. In order to make future losses and benefits associated with each alternative comparable, the standard approach is to convert them into their present values by attaching some weight to each period. This procedure is known as *discounting*, and the weights are called *discount factors*.

Intertemporal choice is an important area of Behavioural Economics and Decision Theory. The are three main interrelated themes of the thesis:

- Axiomatization of time preferences,
- Aggregation of time preferences,
- The effect of delay on the ranking of sequences of dated outcomes.

1.1 Background

In this section we define some notation which will be used throughout the thesis. We discuss different types of impatience and some specifications of discount functions.

We also provide a brief overview of the existing frameworks for axiomatization of time preferences. Some results on aggregation of time preferences, motivated by group decision-making or by some uncertainty about the discount function, will also be given.

1.1.1 Notation

To define the key concepts in the area we need to introduce some notation.

Consider preferences over sequences of dated outcomes. Points in time are elements of the set $T = [0, \infty)$. When we consider a continuous time framework, $t = 0$ is associated with a present moment. Normally the set of outcomes will be assumed to be the interval $X = [0, \infty)$. As in Harvey [43], we treat $x = 0$ as a "neutral", or status quo, outcome. The set of outcomes will be re-defined in some chapters when required.

Let $\mathcal{A}_n = \{\ (\mathbf{x}, \mathbf{t}) \in X^n \times T^n \mid t_1 < t_2 < \ldots < t_n\ \}$ be the set of sequences with n dated outcomes. Define the set of alternatives $\mathcal{A}$ as follows: $\mathcal{A} = \cup_{n=1}^{\infty} \mathcal{A}_n$. Elements of $\mathcal{A}_1 \subseteq \mathcal{A}$ are called *dated outcomes*. Dated outcomes will be denoted (x, t) rather than $((x), (t))$.

For given $\lambda > 0$ and given n define $\mathcal{D}_n^\lambda \subset \mathcal{A}_n$ such that $t_{i+1} - t_i = t_1 = \lambda$ for all $i \in \{1, 2, \ldots, n-1\}$. For given $\lambda > 0$ define

$$\mathcal{D}_\infty^\lambda = \{\ (\mathbf{x}, \mathbf{t}) \in X^\infty \times T^\infty \mid t_{i+1} - t_i = t_1 = \lambda \ \text{ for all } i \geq 1\}.$$

These represent discrete-time sequences (finite or infinite) with fixed period length, λ. Note that $\mathcal{D}_n^\lambda \subseteq \mathcal{A}_n$ but $\mathcal{D}_\infty^\lambda$ is not contained in $\mathcal{A}$, since the latter contains only finite sequences, though of arbitrary length and horizon. In this discrete setting, the sequences will be given a different interpretation. It will be assumed that outcomes are constant within periods: outcome x_i is received throughout the period $[t_{i-1}, t_i)$. In this Introduction we use τ to denote the period $[t_{\tau-1}, t_\tau)$ that ends at time t_τ. (Note that the "present" period corresponds to $\tau = 1$). In this discrete setting, the discount function will be expressed as a function of the period index, τ. In particular, the meaning of $D(\tau)$ for period τ in a discrete setting is different to the interpretation of $D(t)$ in a continuous setting, even when $t = \tau$.

Within subsequent chapters, time will either be continuous or discrete. In the latter case, we will usually suppress the period length (λ) in notation, and use t rather than τ as the period index, since no ambiguity will arise. Following standard convention, we will also simplify notation for discrete-time sequences by treating them as elements of X^n or X^∞ rather than $\mathcal{D}_n^\lambda$ or $\mathcal{D}_\infty^\lambda$ (respectively).

Consider a preference order $\succcurlyeq$ on the set of alternatives X^n for some $n \in \{1, 2, ..., \infty\}$. We say that U is a *discounted utility (DU) representation* for $\succcurlyeq$ if U represents $\succcurlyeq$ and there exist (u, D), such that $u\colon X \to \mathbb{R}$ is a utility function (continuous, strictly increasing, $u(0) = 0$), $D\colon \mathbb{N} \to (0, 1]$ is a discount function (strictly decreasing, $D(1) = 1$ and $\lim_{\tau\to\infty} D(\tau) = 0$) and

$$U(\mathbf{x}) = \sum_{\tau=1}^{n} D(\tau)u(x_\tau)$$

for every $\mathbf{x} \in X^n$. The axiomatization of this representation when the sequences of outcomes differ only in a finite number of periods can be found in [29]. It is useful to note that there exists an axiomatization for X^∞ when X has more structure — as will be explained in Chapter 2. There also exist axiomatizations for X^∞ when more structure is imposed on the discount function, as will be discussed in Section 1.1.4.

Next, consider a preference order $\succcurlyeq$ on the set of alternatives $\mathcal{A}$.

We say that U is a *discounted utility (DU) representation* for $\succcurlyeq$ if U represents $\succcurlyeq$ and there exist (u, D), such that $u\colon X \to \mathbb{R}$ is a utility function (continuous, strictly increasing, $u(0) = 0$), $D\colon T \to (0, 1]$ is a discount function (strictly decreasing, $D(0) = 1$ and $\lim_{t\to\infty} D(t) = 0$) and

$$U(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} D(t_i)u(x_i)$$

for all n and every $(\mathbf{x}, \mathbf{t}) \in \mathcal{A}_n$.

Necessary and sufficient conditions for the existence of a DU representation for preferences on the set $\mathcal{A}$ were provided by Harvey [43, Theorem 2.1]. Fishburn and Rubinstein [32] provided an axiomatization of a DU representation when preferences are restricted to $\mathcal{A}_1$. More information on representations will be given in Section 1.1.4.

1.1.2 Types of impatience and preference reversals

In this section we consider different types of impatience (or aversion to delay).

A property frequently attributed to time preferences is the following:

Definition 1.1 ([63]). *We say that $\succcurlyeq$ exhibits* stationarity, or constant impatience, *if for all $(x, t), (y, s) \in \mathcal{A}_1$ such that $0 < x < y$ and for all $t < s$, $(x, t) \sim (y, s)$ implies $(x, t + \sigma) \sim (y, s + \sigma)$ for any $\sigma > 0$.*

In other words, *constant impatience*, or *stationarity* of preferences means that delaying two dated outcomes by the same amount of time does not change the preferences of a decision-maker between them. When preferences have a DU representation, stationarity restricts only the discount function. The next proposition follows directly from the definition:

Proposition 1.2. *Suppose that* $\succcurlyeq$ *restricted to* $\mathcal{A}_1$ *has a DU representation. Then* $\succcurlyeq$ *exhibits constant impatience if and only if*

$$\frac{D(t+\sigma)}{D(t)} = \frac{D(s+\sigma)}{D(s)}, \text{ for all } t, s \text{ such that } t < s, \text{ and every } \sigma > 0. \tag{1.1}$$

The *discount factor between* $t+\sigma$ *and* t is the ratio $\frac{D(t+\sigma)}{D(t)}$. This factor is constant in t for stationary preferences. It is known that the only discount function that exhibits stationarity is exponential discounting. The exponential model for time discounting has the following form:

$$D(t) = e^{-rt} = \delta^t, \text{ where } r > 0, \delta = e^{-r} \text{ and } \delta \in (0,1).$$

This implicitly defines the meaning of δ^t for irrational t values. That is, $\delta^t = \exp(t \ln \delta)$. We impute the same meaning to δ^t throughout. In a discrete time framework (where $\tau = 1$ corresponds to present) this discount function has the form

$$D(\tau) = \delta^{\tau-1}, \text{ where } \delta \in (0,1).$$

Exponential discounting was introduced in 1937 by Samuelson [72] and still remains the most widely used. The exponential model for time discounting has a well-established axiomatic foundation due to Koopmans [49], [51], with a recent refinement in [17].

However, there has been much experimental research done in recent years that has revealed that decision-makers often tend to violate constant impatience. An excellent review of the variety of time preferences and respective discount models was conducted by Frederick et al. in [33]. These violations of constant impatience can be roughly divided into two groups: those exhibiting decreasing impatience and those exhibiting increasing impatience. Consider the following notions of decreasing and increasing impatience.

Definition 1.3 ([63], [77])**.** *We say that* $\succcurlyeq$ *exhibits* [strictly] decreasing impatience*, if for all* $(x,t), (y,s) \in \mathcal{A}_1$ *such that* $0 < x < y$ *and for all* $t < s$*,* $(x,t) \sim (y,s)$ *implies*

$$(x, t+\sigma) \; [\prec] \preccurlyeq (y, s+\sigma) \tag{1.2}$$

for any $\sigma > 0$*. We say that* $\succcurlyeq$ *exhibits* [strictly] increasing impatience*, if the preference in* (1.2) *is reversed.*[1]

When preferences have a DU representation, decreasing impatience (DI) and in-

[1] Prelec [63] introduced the notions of decreasing and strictly decreasing impatience, but not [strictly] increasing impatience. The definition of *strictly* increasing impatience given here corresponds to Takeuchi's definition of increasing impatience [77].

creasing impatience (II) restrict only the discount function.

Proposition 1.4. *Suppose that $\succcurlyeq$ restricted to $\mathcal{A}_1$ has a DU representation. Then $\succcurlyeq$ exhibits [strictly] DI if and only if*

$$\frac{D(t+\sigma)}{D(t)} \; [<] \leq \frac{D(s+\sigma)}{D(s)}, \; \text{for all } t, s \text{ such that } t < s, \text{ and every } \sigma > 0. \tag{1.3}$$

Furthermore, $\succcurlyeq$ exhibits [strictly] II if and only if the inequality in (1.3) *is reversed.*

DI can be interpreted as follows: if a decision-maker is indifferent between a smaller-sooner reward and a larger-later reward, then when both rewards are delayed (shifted to the future) by the same amount of time, the larger-later reward is preferred to the smaller-sooner reward. DI is the most typical observation in the experiments reviewed by Frederick et al. [33].[2] A recent study by Bleichrodt et al. [16] measured deviations from exponential discounting across money and health domains. The subjects of their experiments exhibited heterogeneity of time preferences with the majority of the participants characterised as decreasingly impatient.

While DI has been most often reported in the experimental literature, there is also some evidence for II (e.g., [11], [77], [74]). Increasing impatience is observed if, whenever the decision-maker is indifferent between a smaller-sooner outcome and a larger-later outcome, the smaller-sooner outcome is preferred if both outcomes are delayed by a common amount. The evidence for II is exemplified in the work undertaken by Attema et al. [11], where it was found that subjects are increasingly impatient for short delays and constantly impatient for longer delays. As Attema et al. conclude: "there is more increasing impatience than commonly believed". In the above-mentioned studies by Bleichrodt et al. [16], the proportion of subjects who displayed increasing impatience was far from being negligible with a range from 25% to 35%. The results of experiments conducted by Takeuchi [77] are even more striking, as it was shown that strictly II was observed in significantly more cases than strictly DI. Of all the subjects in his experiment, around 66% of subjects exhibit strictly II, whereas only 17% are consistent with strictly DI. Takeuchi [77] reported that strictly II is mainly found within a short period of time (around 22 days from the present moment), whereas strictly DI is observed after that period. Evidence of a similar time preference pattern was presented in [74], where subjects exhibited II within one week from the present moment, but were decreasingly impatient within longer periods.

Some explanations for why II is rarely revealed in experimental studies were proposed by Takeuchi [77]. Takeuchi argues that the reason why strictly DI is typically

[2] While DI and II are global properties, experimental evidence can establish consistency with DI and II properties only over a limited range of time periods and outcomes. Hence, experimental results may be described, somewhat loosely, as implying DI over some range and II over another.

reported might be the design of the experiments; i.e., the length of the delay considered. Since previous experimental research has been mainly focused on time preferences within long time delays, it could not elicit strictly II, which occurs within short delays ([74], [77]).

Another property frequently observed in experiments is *present bias*. While various definitions of present bias can be found in the literature (see, for example, [41], [57], [46])[3], the essence of this term can be illustrated by the following example: a decision-maker prefers \$100 today to \$115 tomorrow, but also prefers \$115 in a year and a day to \$100 in a year [33].[4] Note that the second pair of alternatives coincides with the first pair delayed by one year. Present bias means that shifting two dated outcomes - one of which is dated at the present time and another one at some future time - an equal distance into the future may result in a reversal of preferences between these two dated outcomes: from a preference for the earlier outcome to preference for the later one. *Strong present bias* means that delaying two different dated-outcomes by the same amount of time can reverse the ranking of these outcomes from a preference for the sooner to a preference for the later outcome. We call this condition *strong* present bias because it says that a preference reversal may occur even if the earlier of the compared outcomes is dated after the present time. It reflects the general idea that the closer the dated outcomes are to the present the greater the relative attractiveness of the earlier outcome. It will be shown in Chapter 3 that when preferences have a DU representation, the notions of strong present bias and strictly DI are equivalent. Obviously, strong present bias implies present bias. *(Strong) future bias* is defined analogously to (strong) present bias, with the preferences reversed.

Both present bias and future bias illustrate a *reversal of preferences* induced by delay. It is implicitly assumed that as delay increases no more than one reversal can occur [56]. This assumption is similar in essence to the *one-switch property* introduced by Bell [12]. The one-switch property was initially formulated for preferences over lotteries [12]. It says that the preference ranking of any two lotteries is either independent of wealth, or else there is a unique level of wealth such that one strict ranking prevails for lower wealth levels and the opposite strict ranking for higher wealth levels. For preferences with an expected utility representation, the one-switch property restricts the form of the Bernoulli utility function. As demonstrated by Bell [12], the utility functions that satisfy this property are the quadratic, the sum of exponentials, the linear plus exponential and the linear times exponential. The properties of these functions, and their possible applications, have been extensively investigated in risk

[3] Takeuchi's [77] definition of present bias and Takeuchi's [77] definition of DI are equivalent. Analogously, Takeuchi's [77] definition of future bias is equivalent to Takeuchi's [77] definition of II.

[4] Inconsistency in the use of the terms present-bias, strong present bias and (strictly) decreasing impatience in the existing literature will be further discussed in Chapter 3.

theory (see, for example, [1], [2], [14] and [15]). However, it is less well known that Bell [12] also defined an analogous one-switch property for preferences over sequences of dated outcomes. In this case, the one-switch property concerns the effect of adding a common delay to two sequences of dated outcomes: it says that the preference ranking of the delayed sequences is either independent of the delay, or else there is a unique delay such that one strict ranking prevails for shorter delays and the opposite strict ranking for longer delays. Bell [12] claims that if preferences have a DU representation, then the only discount functions consistent with the one-switch property are sums of exponentials.

Collectively, these studies outline the rationale for identifying discount functions capable of accommodating different types of impatience and restrictions on preference reversals. Much effort has been expended to develop models of discounting with better descriptive accuracy than the exponential benchmark. In his recent study Doyle [24] surveys more than twenty models of time discounting, developed by psychologists and economists. In the following section we will consider only the discount functions that will be used in this thesis.

1.1.3 Specifications of discount functions

Since preferences that exhibit DI and present bias are frequently observed in experiments, a number of discount functions have been developed to accommodate these phenomena. One such class of models, favoured by many applied researchers, is *hyperbolic discounting*. Several possible hyperbolic discount functions have been introduced including quasi-hyperbolic discounting ([60], [52]), proportional hyperbolic discounting ([45], [43]) and generalized hyperbolic discounting ([53], [4]).

Let us briefly consider the key types of hyperbolic discounting.

Quasi-hyperbolic discounting is typically defined for discrete-time settings, and is characterised by minimal departure from exponential discounting:

$$D(\tau) = \begin{cases} 1 & \text{if } \tau = 1, \\ \beta\delta^{\tau-1} & \text{if } \tau \geq 2. \end{cases}$$

for some $\delta \in (0, 1)$ and $\beta \in (0, 1]$. Note that when $\beta = 1$ quasi-hyperbolic discounting coincides with exponential discounting. Overall, β defines the degree of deviation from stationarity and can be interpreted as a measure of subjective distance between now ($\tau = 1$) and the immediate future ($\tau = 2$). Quasi-hyperbolic discounting exhibits present bias.

The results of the recent experiments by Chark et al. [21] suggest that decision-makers may be decreasingly impatient within the near future, but discount the remote future (from some period T) at a constant rate.[5] This motivates a further generalization of quasi-hyperbolic discounting, which we call *semi-hyperbolic discounting*:

$$D(\tau) = \begin{cases} 1 & \text{if } \tau = 1, \\ \prod_{i=1}^{\tau-1} \beta_i \delta & \text{if } 1 < \tau \leq T, \\ \delta^{\tau-T} \prod_{i=1}^{T-1} \beta_i \delta & \text{if } \tau > T, \end{cases}$$

where $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_{T-1}$, and $\beta_\tau \in (0, 1]$ for all $\tau \leq T - 1$ and $\delta \in (0, 1)$. We use $SH(T)$ to denote this discount function (for given $\delta, \beta_1, \ldots, \beta_{T-1}$). This form of discounting was previously applied to model the time preferences of a decision-maker in a consumption-savings problem [83]. Our $SH(T)$ specification is not quite the same as the notion of semi-hyperbolic discounting used in [58]. They apply the term to any discount function which satisfies $D(\tau) = \delta^{\tau-T} D(T)$ for all $\tau > T$ (for some T). This class includes $SH(T)$, but is wider. For example, it would allow preferences to exhibit II over the "extended present", as observed by Takeuchi [77]. The possibility of generalizing quasi-hyperbolic discounting was earlier suggested by Hayashi [44]. The form of the discount function he proposed is:

$$D(\tau) = \begin{cases} 1 & \text{if } \tau = 1, \\ \prod_{i=1}^{\tau-1} \beta'_i & \text{if } 1 < \tau \leq T, \\ \delta^{\tau-T} \prod_{i=1}^{T-1} \beta'_i & \text{if } \tau > T, \end{cases}$$

where $\delta \in (0, 1)$ and $0 < \beta'_1 \leq \beta'_2 \leq \ldots \leq \beta'_{T-1} \leq \delta$. By substituting $\delta\beta_\tau = \beta'_\tau$ for all $\tau \leq T - 1$ it is not difficult to see that semi-hyperbolic discounting $SH(T)$ coincides with the form suggested by Hayashi [44]. It is worth mentioning that Hayashi [44] does not provide an axiomatization of this form of discounting, pointing out that this case is somewhat complicated. An axiomatization of semi-hyperbolic discounting will be given in Chapter 2.

The evidence of Chark et al. [21] on extended present bias suggests the following restrictions on the coefficients in $SH(T)$: $\beta_1 < \beta_2 < \ldots < \beta_{T-1}$. In our version of $SH(T)$ we will impose the weaker requirements $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_{T-1}$, and $\beta_\tau \in (0, 1]$ for all $\tau \leq T - 1$ and $\delta \in (0, 1)$. Imposing these restrictions gives some advantages,

[5] Note that these findings are inconsistent with the earlier discussed experimental results of Takeuchi [77].

as it can be immediately seen that exponential and quasi-hyperbolic discounting are the special cases of semi-hyperbolic discounting: $SH(1)$ is the exponential discount function, whereas $SH(2)$ is the quasi-hyperbolic discount function.

Finally, another possible generalization of quasi-hyperbolic discounting for continuous time was offered by Pan et al. [59]. The discount function they use is called *Two-Stage Exponential (TSE) discounting*:

$$D(t) = \begin{cases} \alpha^t, & \text{if } t \leq \lambda, \\ (\frac{\alpha}{\beta})^\lambda \beta^t, & \text{if } t > \lambda, \end{cases}$$

where $t \in [0, T]$, $\alpha, \beta \in [0, 1]$, and $\lambda \in [0, T]$ is called a switch point. The key characteristics of TSE discounting are that it has a constant discount factor α before the switch point λ and a constant discount factor β after this switch point. The coefficient $(\frac{\alpha}{\beta})^\lambda$ is included to guarantee the continuity of the discount function. Pan et al.'s [59] axiomatization of TSE is restricted to single dated outcomes. Although TSE discounting is given for continuous time, it can be viewed as an alternative generalization of quasi-hyperbolic discounting and is distinctively different from semi-hyperbolic discounting.

While quasi-hyperbolic and semi-hyperbolic discount functions partially preserve the exponential form, the next group of discount functions – the generalized hyperbolic and proportional hyperbolic discount functions – have very different functional forms. *Generalized hyperbolic discounting* [53] is described by the following function (for a continuous time environment):

$$D(t) = (1 + ht)^{-\alpha/h}, \text{ where } \alpha, h > 0.$$

Coefficient h measures the level of departure from exponential discounting. As h goes to zero, the discount function becomes exponential $D(t) = e^{-\alpha t}$. Generalized hyperbolic discounting exhibits strictly DI. A special case of generalized hyperbolic discounting when $\alpha = h$ is called *proportional hyperbolic discounting* ([43], [45]), and has a functional form:

$$D(t) = (1 + ht)^{-1}, \text{ where } h > 0.$$

The advantage of this discount function, in comparison with generalized hyperbolic discounting, is that it has only one parameter.

Finally, we must mention some other discount functions that elaborate the exponential form but are less well-known than the quasi-hyperbolic or semi-hyperbolic functions.

Read [68] suggested the use of *subadditive discount functions*. The term "subadditivity" is used by Read to refer to the situation when discounting over a given delay is less when the delay is not divided into intervals than when the delay is divided. In his experimental study, Read [68] derived the discount factors based on the subjects' indifference between smaller-sooner and larger-later rewards, which was obtained by varying the value of the reward. A 24 month delay was divided into three intervals: 0-8 months, 8-16 months, and 16-24 months. For each of three intervals, an indifference between a smaller reward at the beginning of the interval and a bigger reward at the end of the interval was experimentally identified. An indifference was also derived over the undivided delay of 24 months, that is, between a smaller reward in 0 months (now) and a bigger reward in 24 months. Under an assumption of DU representation, it is possible to impute the implied discount factors for each interval. Multiplying the discount factors of three 8-month intervals and comparing the result to the discount factor over the undivided 24-month delay gave evidence for subadditivity of discounting. Read [68] claims that subadditivity can account for DI. He suggested the use of the following variation on exponential discounting for discrete time:

$$D(\tau) = \delta^{\tau^s},$$

where s describes non-linear time perception. Jamison and Jamison [48] refer to a continuous version of Read's model as a *slow Weibul discount function*:

$$D(t) = e^{-rt^{\alpha}},$$

where α indicates non-linear time preference. A similar form of discounting was introduced by Ebert and Prelec [25]. The discount function

$$D(t) = e^{-(rt)^s} \text{ with } r, s > 0$$

is called a *constant-sensitivity* [25] discount function, where the s-parameter allows the function to accommodate strictly DI ($s < 1$) or strictly II ($s > 1$). Clearly, constant-sensitivity discounting is equivalent to exponential discounting when $s = 1$.

McClure et al. [54] studied *double exponential* discounting. They used neuroimaging to analyse the brain areas activated in the process of making intertemporal choice. Two neural systems with different levels of impatience were identified. Based on this, a (discrete time) double exponential discounting model was proposed where each system is represented by exponential discounting as follows:

$$D(\tau) = \lambda\alpha^{\tau-1} + (1-\lambda)\delta^{\tau-1}, \text{ where } \lambda, \alpha, \delta \in (0,1). \tag{1.4}$$

Note also that double exponential discounting can include quasi-hyperbolic discounting by letting $\delta = 0$ in (1.4). This was observed by Xue [82, Proposition 1]).

The respective continuous version of double-exponential discounting is:

$$D(t) = \lambda e^{-rt} + (1-\lambda)e^{-st}, \text{ where } \lambda \in (0,1) \text{ and } r, s > 0.$$

Double-exponential discounting and constant sensitivity discounting are found to be of significant descriptive value [20]. In their recent experimental study, Cavagnaro et al. [20] tested six most common discounting models: exponential, proportional hyperbolic, generalized hyperbolic, quasi-hyperbolic, constant-sensitivity and double exponential discounting. According to their results, constant-sensitivity and double-exponential discounting provide the best fit for the largest number of subjects. Cavagnaro et al. [20] conclude that the greater complexity of these two models is justified by their greater flexibility and that studying the characteristics of these two types of discounting is a promising direction for future research.

Remarkably, double exponential discounting can be considered as a special case of Bell's [12] *sum of exponentials* discounting, which was derived in the context of the one-switch property:

$$D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}, \text{ where } a, b, c > 0, a \leq b/c + 1.$$

Bell's [12] model is more general since it can accommodate strictly II ($1 < a \leq b/c+1$), strictly DI ($a < 1$) and stationarity ($a = 1$), while double exponential discounting with the given parameter restrictions can describe only strictly DI ($\lambda \in (0,1), r \neq s$) and stationarity ($\lambda = 1$).

1.1.4 Existing frameworks and axiomatization of time preferences

The axiomatic foundation of intertemporal decisions is a fundamental question in economics and generates considerable research interest. Despite the fact that a number of models of discounting have appeared in the literature [24], two types have been predominantly used: exponential discounting, first introduced by Samuelson [72], and quasi-hyperbolic discounting [60, 52]. The important question to be answered is which axioms allow us to say that the preferences of a decision-maker can be represented using the DU model with exponential or quasi-hyperbolic discount functions? Existing axiom systems for intertemporal decisions address this question. These systems can be roughly divided into two main groups: those with preferences over deterministic consumption streams and those with preferences over stochastic consumption streams.

The first group has been the leading approach in the area, both for exponential and quasi-hyperbolic functions. In this framework a consumption set is endowed with topological structure, and Debreu's [23] theorem on additive representation is a key mathematical tool.

Koopmans' result for exponential discounting in discrete time with deterministic consumption streams [49, 50, 51] remains the best known. A revised formulation of Koopmans' result was proposed by Bleichrodt et al. [17], using alternative conditions on preferences. A similar approach was also suggested by Harvey [42], who works in continuous time. The axiomatic foundation of exponential discounting for the special case of preferences over $\mathcal{A}_1$ (single dated outcomes) was presented by Fishburn and Rubinstein [32]. Their axioms will often be assumed in following chapters, so we state them here for future reference. The list of the axioms is as follows:

Axiom 1. (Weak Order) The preference order $\succcurlyeq$ is a weak order; i.e., it is complete and transitive.

Axiom 2. (Monotonicity) For every $x, y \in X$, if $x < y$, then $(x,t) \prec (y,t)$ for every $t \in T$.

Axiom 3. (Continuity) For every $(y,s) \in \mathcal{A}_1$ the sets $\{(x,t) \in \mathcal{A}_1 \mid (x,t) \succcurlyeq (y,s)\}$ and $\{(x,t) \in \mathcal{A}_1 \mid (x,t) \preccurlyeq (y,s)\}$ are closed.

Axiom 4. (Impatience) For all $t, s \in T$ and every $x > 0$, if $t < s$, then $(x,t) \succ (x,s)$. If $t < s$ and $x = 0$, then $(x,t) \sim (x,s)$ for every $t, s \in T$; that is, 0 is a time-neutral outcome.

Axiom 5. (Separability) For every $x, y, z \in X$ and every $r, s, t \in T$ if $(x,t) \sim (y,s)$ and $(y,r) \sim (z,t)$ then $(x,r) \sim (z,s)$.

Fishburn and Rubinstein [32] proved the following result:

Theorem 1.5 ([32])**.** *The preferences $\succcurlyeq$ on $\mathcal{A}_1$ satisfy Axioms 1-5 if and only if there exists a DU representation for $\succcurlyeq$ on $\mathcal{A}_1$. If (u, D) and (u_0, D_0) both provide DU representations for $\succcurlyeq$ on $\mathcal{A}_1$, then $u = \alpha u_0$ for some $\alpha > 0$, and $D = \beta D_0$ for some $\beta > 0$.*

In a non-stochastic framework with discrete time, quasi-hyperbolic discounting has been axiomatized by Olea and Strzalecki [58], building on Bleichrodt et al. [17]. Pan et al. [59] provide an axiomatization of their continuous-time generalization of quasi-hyperbolic discounting (two-stage exponential discounting) for preferences over single dated outcomes.

All the axiomatization systems mentioned above are formulated for consumption streams over an infinite horizon; i.e., for $\mathcal{A}$ or $\mathcal{D}_\infty^\lambda$. The finite horizon case has rarely been discussed. For exponential discounting in a discrete time setting, however, it can be found in [30].

The second group of axiomatic systems considers stochastic consumption streams in discrete time. These are lotteries with outcomes in $\mathcal{D}_\infty^\lambda$ for some given λ. To obtain an additive form the fundamental representation theorem of von Neumann and Morgenstern (vNM) [78] is used. The application of this approach to exponential discounting was given by Epstein [26]. The axiomatization of quasi-hyperbolic discounting by Hayashi [44] builds on Epstein's axiom system [26].

1.1.5 Time preferences: aggregation and uncertainty

Sometimes decisions about dated outcomes have to be made by a group of individuals, such as boards, committees or households. It is natural to think that individuals may differ in the discounting procedure they use. If the decision is to be made by the group it is desirable to have an aggregating procedure that suitably reflects the time preferences of all members. The natural option is to average the discount functions across individuals, which is equivalent to averaging the discounted utilities in the case when all agents have identical utility functions. This approach guarantees that the aggregation satisfies Pareto's principle and has been widely used in the existing literature on time preferences (for example, see [46], [84], [82]). It is known that such collective discount functions need not share properties that are common to the individual discount functions being aggregated. As Jackson and Yariv demonstrate [46], if individuals discount the future exponentially and there is a heterogeneity in discount factors, then their aggregate discount function exhibits strong present bias. Moreover, when the number of individuals grows, in the limit the group discount function becomes hyperbolic [46].

Jackson and Yariv [46] give the following example of present-biased group preferences for a household with two individuals, Constantine and Patience. Both have identical instantaneous utility functions, and discount the future exponentially, but Constantine has a discount factor of 0.5, whereas Patience has a discount factor of 0.8. Suppose that they need to choose between 10 utiles for each today or 15 utiles for each tomorrow. They calculate the aggregate discounted utility for each option: $10 + 10 = 20$ and $15(0.8 + 0.5) = 19.5$. Therefore, 10 utiles today is chosen. Now suppose that they must choose between 10 utiles at time $t \geq 1$ and 15 utiles at $t + 1$. The aggregate discounted utilities in this case are $10(0.8^t + 0.5^t)$ and $15(0.8^{t+1} + 0.5^{t+1})$, respectively. For any $t \geq 1$ the 15 utiles at $t + 1$ is preferable to the 10 utiles at t, which reverses the initial preference for 10 utiles at $t = 0$ over 15 utiles at $t = 1$. The behaviour of the household is present-biased. Jackson and Yariv [46] show that this household in fact exhibits *strong* present bias.

A related result is obtained by Prelec [63]. Prelec [63] introduces a natural notion of comparative DI and proves that one individual exhibits more DI preferences than another if the logarithm of the discount function of the former is more convex than that of the latter. Using this notion of comparative DI, Prelec [63] demonstrates that the mixture of two distinct discount functions of equally DI preferences, represents preferences which are more DI.

Another scenario in which the aggregation of time preferences may be required is when a single decision-maker is uncertain about the appropriate discount function to apply. For example, discounting may be affected by a survival function with a constant but uncertain hazard rate. Such scenarios are considered by Weitzman [81] and Sozou [75].[6] If the decision-maker maximizes expected discounted utility, then she maximizes discounted utility for a certainty equivalent discount function, calculated as the probability-weighted average of the different possible discount functions that may apply. Weitzman [81] shows that if each of the possible rates of time preference converges to some non-negative value (as time goes to infinity), then the certainty equivalent time preference function converges to the lowest of these limits. Weitzman's result implies that long-term cost and benefits should be discounted at the largest of the possible discount factors. This corresponds to the discount function which, as time goes to infinity, represents the most patient time preferences. Working in a similar framework, Sozou [75] considers a decision-maker whose discounting reflects a survival function with a constant, but uncertain, hazard rate. If this hazard rate is exponentially distributed, Sozou shows that the decision-maker's expected discount function is hyperbolic.

1.2 Outline of thesis

In Chapter 2 we provide a new and simple axiomatization of exponential and quasi-hyperbolic discounting. We work with preferences over streams of consumption *lotteries*; i.e., a setting in which there is a lottery in each period of time. In other words, we restrict Epstein and Hayashi's framework ([26], [44]) to product measures. This framework allows us to apply Anscombe and Aumann's result from Subjective Expected Utility Theory [8] and to obtain a discounted *expected* utility (DEU) representation

[6] Farmer and Geanakoplos [28] analyse the behaviour of a certainty equivalent discount function as time goes to infinity when time preference rates are stochastic and change according to geometric random walk. Geanakoplos et al. [34] investigate the same problem for a more general class of random walks.

using relatively simple axioms. Importantly, we establish a unified treatment of exponential and quasi-hyperbolic discounting in both finite and infinite settings. With Fishburn [31] and Harvey [42] as the key sources of technical inspiration, our approach also facilitates proofs that are relatively straightforward. In addition, we provide an axiomatization of semi-hyperbolic discounting, which, in our framework, can be obtained using a simple elaboration of the axioms for quasi-hyperbolic discounting.

In Chapter 3 we consider the aggregation of preferences over single dated outcomes which exhibit decreasing impatience. The goal of this chapter is twofold. Firstly, we seek to extend the results of Prelec [63] and Jackson and Yariv [46] on the aggregation of discount functions. We show that when preferences satisfy the axioms of Fishburn and Rubinstein [32], strictly decreasing impatience (or strictly increasing discount factor) is equivalent to strong present bias. Theorem 3.20 establishes that the weighted average of discount functions that are from the same DI class (in the sense of Prelec [63]) always exhibits *strictly* more DI than each component. This generalizes both Prelec's [63] and Jackson and Yariv's [46] result.

The second goal of Chapter 3 is to prove an analogue of Weitzman's [81] result: one in which discounting is hyperbolic rather than exponential, but there is uncertainty about the hyperbolic discount factor. Our result, given in Theorem 3.28, is very different to Weitzman's. We show that the certainty equivalent hyperbolic discount factor converges, not to the lowest individual hyperbolic discount factor, but to the probability-weighted harmonic mean of the individual hyperbolic discount factors.

In Chapter 4 we analyse the one-switch property for intertemporal preferences introduced by Bell [12]. For preferences that have a DU representation, Bell [12] argues that the only discount functions consistent with the one-switch property are sums of exponentials. This chapter proves that discount functions of the linear times exponential form also satisfy the one-switch property. We further demonstrate that preferences which have a DU representation with the linear times exponential discount function exhibit increasing impatience ([77]). We also clarify an ambiguity in the original Bell [12] definition of the one-switch property by distinguishing a weak one-switch property from the (strong) one-switch property. We show that the one-switch property and the weak one-switch property definitions are equivalent in a continuous-time version of the Anscombe and Aumann [8] setting.

Finally, in Chapter 5 we offer some conclusions and discuss possible directions for future research.

2

A simple framework for the axiomatization of exponential and quasi-hyperbolic discounting

The results in this chapter have been published in [6].

In the previous chapter we introduced exponential, quasi-hyperbolic and semi-hyperbolic discount functions. We also provided a brief survey of existing axiomatization systems for these types of discounting (Section 1.1.4). The goal of this chapter is to suggest an alternative way to axiomatize exponential, quasi-hyperbolic and semi-hyperbolic discounting. Note that semi-hyperbolic discounting subsumes quasi-hyperbolic discounting as a special case. The chapter is organized as follows. We first give preliminaries and define the setting. The Anscombe and Aumann (AA) approach uses "horse lotteries" (with unknown probabilities) and "roulette lotteries" (with known probabilities). In the AA environment preferences are expressed over compounds of "horse" and "roulette" lotteries: each "horse" is associated with a particular "roulette lottery". The outcome of such compound lottery is resolved in two stages. In the first stage the true state of the nature (the winning "horse") is identified. In the second stage the roulette lottery corresponding to the true state of nature is resolved. We recall the Anscombe and Aumann [8] result in Section 2.2.1, and reinterpret it for our intertemporal context by re-labeling states of nature as periods of time. In Section

2.2.2 we consider its extension to an infinite horizon (infinitely many states). Using these results as a foundation for our axiomatization, we first obtain the axiomatization for exponential and semi-hyperbolic discounting for a finite time horizon (Section 2.3). We next consider an infinite time horizon (Section 2.4). We end in Section 2.5 with a discussion of the results.

2.1 Preliminaries

Assume that the objectives of a decision-maker can be expressed by a preference order $\succcurlyeq$ on the set of alternatives X^n, where n may be ∞. Think of these alternatives as discrete-time sequences of outcomes, for time periods $t \in \{1, 2, \dots, n\}$.We say that a utility function $U\colon X^n \to \mathbb{R}$ *represents* this preference order, if for all $\mathbf{x}, \mathbf{y} \in X^n$, $\mathbf{x} \succcurlyeq \mathbf{y}$ if and only if $U(\mathbf{x}) \geq U(\mathbf{y})$.

In this chapter, we assume that X is a *mixture set*. That is, for every $x, y \in X$ and every $\lambda, \mu \in [0, 1]$, there exists $x\lambda y \in X$ satisfying:

- $x1y = x$,
- $x\lambda y = y(1-\lambda)x$,
- $(x\mu y)\lambda y = x(\lambda\mu)y$.

Since X is a mixture set, the set X^n is easily seen to be a mixture set under the following mixture operation: $\mathbf{x}\lambda\mathbf{y} = (x_1\lambda y_1, \dots, x_n\lambda y_n)$, where $\mathbf{x}, \mathbf{y} \in X^n$ and $\lambda \in [0, 1]$.

The utility function $u\colon X \to \mathbb{R}$ is called *mixture linear* if for every $x, y \in X$ we have $u(x\lambda y) = \lambda u(x) + (1-\lambda)u(y)$ for every $\lambda \in [0, 1]$.

The binary relation $\succcurlyeq$ on X^n induces a binary relation (also denoted $\succcurlyeq$) on X in the usual way: for any $x, y \in X$ the preference $x \succcurlyeq y$ holds if and only if $(x, x, \dots, x) \succcurlyeq (y, y, \dots, y)$.

The function U is called a *discounted utility function* if

$$U(\mathbf{x}) = \sum_{t=1}^{n} D(t)u(x_t),$$

for some non-constant $u\colon X \to \mathbb{R}$ and some $D\colon \mathbb{N} \to \mathbb{R}$ with $D(1) = 1$. The function D is called the *discount function*. If u is mixture linear (and non-constant), then the function U is called a *discounted expected utility function*. Note that mixture linear functions are continuous. It is also worth mentioning that we do not impose the requirement that D is strictly decreasing here, but this property will be introduced and axiomatically characterised later.

Recall that there are two types of discount functions which are commonly used in modelling of time preferences:

- *Exponential discounting*: $D(t) = \delta^{t-1}$, where $\delta \in (0,1)$ is called a discount factor.
- *Quasi-hyperbolic discounting*:
$$D(t) = \begin{cases} 1 & \text{if } t = 1, \\ \beta\delta^{t-1} & \text{if } t \geq 2. \end{cases}$$
for some $\delta \in (0,1)$ and $\beta \in (0,1]$.

A further generalization of quasi-hyperbolic discounting is called *semi-hyperbolic discounting*:
$$D(t) = \begin{cases} 1 & \text{if } t = 1, \\ \prod_{i=1}^{t-1} \beta_i \delta & \text{if } 1 < t \leq T, \\ \delta^{t-T} \prod_{i=1}^{T-1} \beta_i \delta & \text{if } \ t > T. \end{cases}$$
We use $SH(T)$ to denote this discount function (for given $\delta, \beta_1, \ldots, \beta_{T-1}$).

2.2 Anscombe and Aumann (1963) representations

A pre-condition for obtaining discounting in an exponential or quasi-hyperbolic form is additive separability. In the framework of preferences over streams of lotteries, Anscombe and Aumann's (1963) theorem [8] provides axioms which give an additively separable representation when $n < \infty$. Anscombe and Aumann formulated their result for acts rather than temporal streams. Here, states of the world are replaced by time periods.

We say that the preference order $\succcurlyeq$ on X^n has an *Anscombe and Aumann (AA) representation*, if for every $\mathbf{x}, \mathbf{y} \in X^n$:
$$\mathbf{x} \succcurlyeq \mathbf{y} \text{ if and only if } \sum_{t=1}^{n} w_t u(x_t) \geq \sum_{t=1}^{n} w_t u(y_t),$$
where $u \colon X \to \mathbb{R}$ is non-constant and mixture linear and w_t are constant weights such that $w_t \geq 0$ for each t with at least one $w_t > 0$. We also say that the pair $(u, \mathbf{w})$ provides an AA representation for $\succcurlyeq$. There are two main differences between a DU representation and an AA representation. First, there is an additional restriction of mixture linearity on utility function in the latter. Second, there are weaker conditions

imposed on the weights w_t in an AA representation than on a discount function in a DU representation: w_t are required to be non-negative with at least one strictly positive w_t, whereas the values of a discount function must be strictly positive at each $t \geq 1$.

2.2.1 Finite case ($n < \infty$)

For $n < \infty$ the following axioms are necessary and sufficient for an AA representation:

Axiom F1. (Weak order). $\succcurlyeq$ is a weak order on X^n.

Axiom F2. (Non-triviality). There exist some $a, b \in X$ such that

$$(a, a, \ldots, a) \succ (b, b, \ldots, b).$$

Axiom F3. (Mixture independence). $\mathbf{x} \succcurlyeq \mathbf{y}$ if and only if $\mathbf{x}\lambda\mathbf{z} \succcurlyeq \mathbf{y}\lambda\mathbf{z}$ for every $\lambda \in (0, 1)$ and every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^n$.

Axiom F4. (Mixture continuity). For every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^n$ the sets $\{\alpha \mid \mathbf{x}\alpha\mathbf{z} \succcurlyeq \mathbf{y}\}$ and $\{\beta \mid \mathbf{y} \succcurlyeq \mathbf{x}\beta\mathbf{z}\}$ are closed subsets of the unit interval.

Axiom F5. (Monotonicity). For every $\mathbf{x}, \mathbf{y} \in X^n$ if $x_t \succcurlyeq y_t$ for every t then $\mathbf{x} \succcurlyeq \mathbf{y}$.

Theorem 2.1 (AA)**.** *The preferences $\succcurlyeq$ on X^n satisfy axioms F1-F5 if and only if there exists an AA representation $(u, \mathbf{w})$ for $\succcurlyeq$ on X^n. Moreover, $(u', \mathbf{w}')$ is another AA representation for $\succcurlyeq$ on X^n if and only if there are some $A > 0$, some B and some $C > 0$ such that $u' = Au + B$ and $\mathbf{w}' = C\mathbf{w}$.*

The proof of the theorem for the general mixture set environment can easily be constructed by combining the arguments in [31] and [71]. Evidently, the key axiom here is the condition of mixture independence. It is a strong axiom which imposes an additive structure.

2.2.2 Infinite case ($n = \infty$)

Anscombe and Aumann's result may be extended to the infinite horizon case. One possible extension is given by Fishburn [31]. However, we give a slightly modified version which incorporates ideas from Harvey [42].

Fix some $x_0 \in X$. We refer to the same x_0 throughout the rest of the chapter. A consumption stream $\mathbf{x}$ is called *ultimately x_0-constant* if there exists T such that $\mathbf{x} = (x_1, \ldots, x_T, x_0, x_0, \ldots)$. Note the difference between this term and the related notion of an "ultimately constant" stream in [17] and [58], which does not fix the value at which consumption is ultimately constant. Let X_T be the set of ultimately x_0-constant consumption streams of length T. Denote the union of the sets X_T over all T as X^*. Let X^{**} be the union of X^* and all constant streams. It is not hard to see that both $X^*, X^{**} \subset X^\infty$ are mixture sets.

We must mention that the fixed x_0 serves two purposes: first, it will be needed to state the convergence axiom; and second, it allows us to define the class X^* of ultimately x_0-constant streams in a way that makes them a strict subset of the usually defined class. Since some of the axioms only restrict preferences over X^{**} this second aspect confers some advantages.

Axiom I1. (Weak order). $\succcurlyeq$ is a weak order on X^∞.

Axiom I2. (Non-triviality). There exist some $a, b \in X$ such that $a \succ x_0 \succ b$.

Axiom I2 implies that x_0 is an interior point with respect to preference. It restricts both $\succcurlyeq$ and the choice of the fixed element x_0. The role of x_0 here is similar to the role of "neutral" outcome $x = 0$ in Chapter 1, since we will later impose that $u(x_0) = 0$. Note, however, that $u(x_0)$ is interior to the range of u.

Axiom I3. (Mixture independence). $\mathbf{x} \succcurlyeq \mathbf{y}$ if and only if $\mathbf{x}\lambda\mathbf{z} \succcurlyeq \mathbf{y}\lambda\mathbf{z}$ for every $\lambda \in (0,1)$ and every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in X^{**}$.

Axiom I4. (Mixture continuity). For every $\mathbf{x}, \mathbf{z} \in X^{**}$ and every $\mathbf{y} \in X^\infty$ the sets $\{\alpha \mid \mathbf{x}\alpha\mathbf{z} \succcurlyeq \mathbf{y}\}$ and $\{\beta \mid \mathbf{y} \succcurlyeq \mathbf{x}\beta\mathbf{z}\}$ are closed subsets of the unit interval.

Axiom I5. (Monotonicity). For every $\mathbf{x}, \mathbf{y} \in X^\infty$: if $x_t \succcurlyeq y_t$ for every t then $\mathbf{x} \succcurlyeq \mathbf{y}$.

We have applied a weaker version of the monotonicity axiom in comparison with the interperiod monotonicity used by Fishburn. However, Axiom I5 is sufficient to obtain an AA representation.

For the statement of the next axiom we need to introduce some notation. Let $[a]_k = (x_0, \ldots, x_0, a, x_0, \ldots)$ where $a \in X$ is in the k^{th} position. Using this notation, we state the following axiom:

Axiom I6. (Convergence). For every $\mathbf{x} = (x_1, x_2, \ldots) \in X^\infty$, every $x^+, x^- \in X$ and every k:

- if $[x^+]_k \succ [x_k]_k$ there exists $T^+ \geq k$ such that

$$\mathbf{x} \preccurlyeq \mathbf{x}^+_{k,T} \text{ for all } T \geq T^+,$$

where $\mathbf{x}^+_{k,T} = (x_1, x_2, \ldots, x_{k-1}, x^+, x_{k+1}, \ldots, x_T, x_0, x_0, \ldots)$;

- if $[x^-]_k \prec [x_k]_k$ there exists $T^- \geq k$ such that

$$\mathbf{x} \succcurlyeq \mathbf{x}^-_{k,T} \text{ for all } T \geq T^-,$$

where $\mathbf{x}^-_{k,T} = (x_1, x_2, \ldots, x_{k-1}, x^-, x_{k+1}, \ldots, x_T, x_0, x_0, \ldots)$.

Our convergence axiom differs from Axiom B6, that was used by Fishburn:

Axiom B6. For some $\hat{x} \in X$, every $\mathbf{x}, \mathbf{y} \in X^\infty$ and every $\lambda \in (0, 1)$:

- if $\mathbf{x} \succ \mathbf{y}$, then there exists T such that $(x_1, \ldots, x_n, \hat{x}, \hat{x}, \ldots) \succcurlyeq \mathbf{x}\lambda\mathbf{y}$ for all $n \geq T$;
- if $\mathbf{x} \prec \mathbf{y}$, then there exists T such that $(x_1, \ldots, x_n, \hat{x}, \hat{x}, \ldots) \preccurlyeq \mathbf{x}\lambda\mathbf{y}$ for all $n \geq T$.

Instead, Axiom I6 adapts ideas from [42].[1] Axiom I6 is more appealing for our purposes as it not only guarantees the convergence of the AA representation, but also allows us to relax two axioms, mixture independence and mixture continuity, which are no longer required to hold on all of X^∞.

We thus obtain the following representation:

Theorem 2.2 (Infinite AA). *The preferences $\succcurlyeq$ on X^∞ satisfy axioms I1-I6 if and only if there exists an AA representation $(u, \mathbf{w})$ for $\succcurlyeq$ on X^∞. Moreover, $(u', \mathbf{w}')$ is another AA representation for $\succcurlyeq$ on X^∞ if and only if there are some $A > 0$, some B and some $C > 0$ such that $u = Au' + B$ and $\mathbf{w} = C\mathbf{w}'$.*

The proof of Theorem 2.2 combines elements of the arguments in [31], [42] and [71].

Proof. Necessity of the axioms is straightforward to verify. Therefore we will focus on the proof of sufficiency.

Step 1. Applying Theorem 1 of [31] to the mixture set X, it follows from Axioms I1, I3, I4 that there exists a mixture linear utility function u preserving the order on X (unique up to positive affine transformations). Normalize u so that $u(x_0) = 0$. Note that by non-triviality $u(x_0)$ is in the interior of the non-degenerate interval $u(X)$.

Convert streams into their utility vectors by replacing the outcomes in each period by their utility values. Define the following order: $(v_1, v_2, \ldots) \succcurlyeq^* (u_1, u_2, \ldots) \Leftrightarrow$ there exist $\mathbf{x}, \mathbf{y} \in X^\infty$ such that $\mathbf{x} \succcurlyeq \mathbf{y}$ and $u(x_t) = v_t$ and $u(y_t) = u_t$ for every t. This order is unambiguously defined because of the monotonicity assumption, i.e., if $x_i \sim x'_i$ then $(x_1, \ldots, x_i, \ldots) \sim (x_1, \ldots, x'_i, \ldots)$.

[1] It is worth mentioning that Fishburn's motivation for the convergence axiom B6 looks somewhat contrived in the context of acts [31, p. 113]. However, it becomes very natural in the context where states of the world are re-interpreted as periods of time.

The preference order $\succcurlyeq^*$ inherits the properties of weak order, mixture independence and mixture continuity from $\succcurlyeq$. Note that $u(X)^\infty$ is a mixture set under the standard operation of taking convex combinations: if $\mathbf{v}, \mathbf{u} \in u(X)^\infty$ then

$$\mathbf{v}\lambda\mathbf{u} = \lambda\mathbf{v} + (1-\lambda)\mathbf{u} \text{ for every } \lambda \in (0,1).$$

Therefore, by Theorem 1 of [31] we obtain a mixture linear representation $U\colon\ u(X)^\infty \to \mathbb{R}$, where U is unique up to positive affine transformations.

Hence $\mathbf{v} \succcurlyeq^* \mathbf{u}$ if and only if $U(\mathbf{v}) \geq U(\mathbf{u})$.

Step 2. Normalize U so that $U(0,0,\ldots) = U(\mathbf{0}) = 0$. Since 0 is in the interior of $u(X)$, and since $U(\mathbf{v}\lambda\mathbf{0}) = \lambda U(\mathbf{v})$ for any $\mathbf{v} \in \mathbb{R}^\infty$ and for every $\lambda \in (0,1)$, we can assume that U is defined on $\mathbb{R}^\infty$.

Mixture linearity of U implies standard linearity of U on $\mathbb{R}^\infty$. To prove this, we need to show that $U(k\mathbf{v}) = kU(\mathbf{v})$ for any k and $U(\mathbf{v}+\mathbf{u}) = U(\mathbf{v}) + U(\mathbf{u})$ for any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^\infty$.

As $u(X)^\infty$ is a mixture set under the operation of taking convex combinations, $U(\mathbf{v}k\mathbf{0}) = U\left(k\mathbf{v} + (1-k)\mathbf{0}\right) = U(k\mathbf{v}) = kU(\mathbf{v})$ for any $k \in (0,1)$. If $k > 1$ then $U(\mathbf{v}) = U\left(\frac{k}{k}\mathbf{v}\right) = \frac{1}{k}U(k\mathbf{v})$. Multiplying both parts of this equation by k, we obtain $U(k\mathbf{v}) = kU(\mathbf{v})$ for all $k > 1$. Therefore, $U(k\mathbf{v}) = kU(\mathbf{v})$ for any $k > 0$.

To prove that $U(\mathbf{v}+\mathbf{u}) = U(\mathbf{v}) + U(\mathbf{u})$, consider the mixture $\mathbf{v}\frac{1}{2}\mathbf{u}$. By mixture linearity of U we have:

$$U\left(\mathbf{v}\frac{1}{2}\mathbf{u}\right) = \frac{1}{2}U(\mathbf{v}) + \frac{1}{2}U(\mathbf{u}) = \frac{1}{2}\left(U(\mathbf{v}) + U(\mathbf{u})\right). \tag{2.1}$$

On the other hand, $\mathbf{v}\frac{1}{2}\mathbf{u} = \frac{1}{2}\mathbf{v} + \frac{1}{2}\mathbf{u} = \frac{1}{2}(\mathbf{v}+\mathbf{u})$. Therefore,

$$U\left(\mathbf{v}\frac{1}{2}\mathbf{u}\right) = U\left(\frac{1}{2}(\mathbf{v}+\mathbf{u})\right) = \frac{1}{2}U(\mathbf{v}+\mathbf{u}) \tag{2.2}$$

Comparing (2.1) and (2.2) we conclude that $U(\mathbf{v}+\mathbf{u}) = U(\mathbf{v}) + U(\mathbf{u})$.

Finally, note that

$$U(\mathbf{0}) = U\left(\mathbf{v} + (-\mathbf{v})\right) = U(\mathbf{v}) + U(-\mathbf{v}) = 0,$$

hence $U(-\mathbf{v}) = -U(\mathbf{v})$. Therefore, if $k < 0$, then $U(k\mathbf{v}) = -kU(-\mathbf{v}) = kU(\mathbf{v})$.

For each T, consider the function $f\colon \mathbb{R}^T \to \mathbb{R}$ defined as follows:

$$f(v_1, \ldots, v_T) = U(v_1, \ldots, v_T, 0, 0, \ldots).$$

This function is linear on $\mathbb{R}^T$ and it satisfies $f(\mathbf{0}) = 0$, therefore,

$$f(v_1, \ldots, v_T) = \sum_{t=1}^{T} w_t^T v_t,$$

where $\mathbf{w}^T = (w_1^T, \ldots, w_T^T)$. By monotonicity $w_t^T \geq 0$ for all $t \leq T$.

Note that $w_t^T = U([1]_t)$, where $[1]_t$ is the vector with 1 in period t and 0 elsewhere. It follows that $w_t^T = w_t^{T'}$ for any T and T'. Hence there is a vector $\mathbf{w} \in \mathbb{R}^\infty$ such that

$$U(v_1, \ldots, v_T, 0, 0, \ldots) = \sum_{t=1}^{\infty} w_t v_t \text{ for any } (v_1, \ldots, v_T) \in \mathbb{R}^T.$$

Recalling that $v_t = u(x_t)$ we obtain

$$U(u(x_1), \ldots, u(x_T), 0, 0, \ldots) = \sum_{t=1}^{T} w_t u(x_t) \text{ for all } \mathbf{x} \in X^*.$$

Therefore, for every $\mathbf{x}, \mathbf{y} \in X^*$ we have $\mathbf{x} \succcurlyeq \mathbf{y}$ if and only if

$$\sum_{t=1}^{T} w_t u(x_t) \geq \sum_{t=1}^{T} w_t u(y_t).$$

By slightly abusing the notation, re-define U so that:

$$U(x_1, \ldots, x_T, x_0, x_0, \ldots) = \sum_{t=1}^{T} w_t u(x_t) \text{ for all } \mathbf{x} \in X^*.$$

Hence $U(\mathbf{x}) = \sum_{t=1}^{\infty} w_t u(x_t)$ represents preferences on X^*.

Step 3. Next, we show that $U(x_1, x_2, \ldots)$ converges for any $(x_1, x_2, \ldots)$. Define $U_T \colon X^\infty \to \mathbb{R}$ as follows:

$$U_T(\mathbf{x}) = \sum_{t=1}^{T} u_t(x_t), \text{ where } u_t(x_t) = w_t u(x_t).$$

Consider the sequence of functions $U_1, U_2, \ldots, U_T, \ldots$ According to the Cauchy Criterion, a sequence of functions $U_T(\mathbf{x})$ defined on X^∞ converges on X^∞ if and only if for any $\varepsilon > 0$ and any $\mathbf{x} \in X^\infty$ there exists $T \in \mathbb{N}$ such that $|U_N(\mathbf{x}) - U_M(\mathbf{x})| < \varepsilon$ for any $N, M \geq T$.

Fix some $\mathbf{x} \in X^\infty$ and $\varepsilon > 0$. Suppose that for some k it is possible to choose x^+, x^- such that $[x^+]_k \succ [x_k]_k \succ [x^-]_k$. By Step 2 the preference $[x^+]_k \succ [x_k]_k \succ [x^-]_k$

implies that $w_k > 0$. Therefore, as u is a continuous function, it is without loss of generality to assume that

$$u_k(x^+) - u_k(x_k) < \varepsilon/2 \text{ and } u_k(x_k) - u_k(x^-) < \varepsilon/2.$$

It follows that $u_k(x^+) - u_k(x^-) < \varepsilon$, or $u_k(x^-) - u_k(x^+) > -\varepsilon$. By Axiom I6 there exist T^+ and T^- satisfying $k \leq \min\{T^-, T^+\}$ such that

$$\mathbf{x}^+_{k,N} \succcurlyeq \mathbf{x} \succcurlyeq \mathbf{x}^-_{k,M}, \text{ for all } N \geq T^+,\ M \geq T^-.$$

Let $T^* = \max\{T^-, T^+\}$. It is necessary to demonstrate that $|U_N(\mathbf{x}) - U_M(\mathbf{x})| < \varepsilon$ for any $N, M \geq T^*$. If $N = M$ the result is obviously true. If $N \neq M$ then it is without loss of generality to assume that $N > M$. By the additive representation:

$$U(\mathbf{x}^+_{k,N}) \geq U(\mathbf{x}^-_{k,M}).$$

Expanding

$$u_k(x^+) + \sum_{t=1,t\neq k}^{N} u_t(x_t) \geq u_k(x^-) + \sum_{t=1,t\neq k}^{M} u_t(x_t).$$

By rearranging this inequality

$$\sum_{t=M+1}^{N} u_t(x_t) \geq u_k(x^-) - u_k(x^+) > -\varepsilon.$$

As $N > M \geq T^*$ it is also true that $U(\mathbf{x}^+_{k,M}) \geq U(\mathbf{x}^-_{k,N})$, hence

$$\sum_{t=M+1}^{N} u_t(x_t) \leq u_k(x^+) - u_k(x^-) < \varepsilon.$$

Note that

$$\sum_{t=M+1}^{N} u_t(x_t) = U_N(\mathbf{x}) - U_M(\mathbf{x}).$$

Hence, $|U_N(\mathbf{x}) - U_M(\mathbf{x})| < \varepsilon$ and it follows that $U(\mathbf{x})$ converges by the Cauchy criterion.

Suppose now that it is not possible to find such k that $[x^+]_k \succ [x_k]_k \succ [x^-]_k$ for some $x^+, x^- \in X$. If $w_t = 0$ for all t then the result is trivial. Suppose that $w_t > 0$ for some t. Then for every period t for which $w_t > 0$ we have

$$x_t \in X^e \equiv \{z \in X \ \mid \ z \succcurlyeq z' \text{ for all } z' \in X \text{ or } z' \succcurlyeq z \text{ for all } z' \in X\}.$$

For some $\lambda \in (0,1)$ replace x_t with the mixture $x_t \lambda x_0$ for each t. Call the resulting stream $\mathbf{x}^*$. Then

$$U_T(\mathbf{x}) - U_T(\mathbf{x}^*) = \sum_{t=1}^{T} u_t(x_t) - \sum_{t=1}^{T} u_t\left(x_t \lambda x_0\right) = (1-\lambda) \sum_{t=1}^{T} u_t(x_t) = (1-\lambda) U_T(\mathbf{x}).$$

By rearranging this equation it follows that $U_T(\mathbf{x}^*) = \lambda U_T(\mathbf{x})$. By the previous argument $U_T(\mathbf{x}^*)$ converges, therefore, $U_T(\mathbf{x})$ converges.

Step 4. Show that $U(\mathbf{x})$ represents the order on X^∞. Suppose that $\mathbf{x} \succcurlyeq \mathbf{y}$, where $\mathbf{x}, \mathbf{y} \in X^\infty$. If for some k, j it is possible to find x^+, y^- such that $[x^+]_k \succ [x_k]_k$ and $[y^-]_j \prec [y_j]_j$, then $[x^+ \lambda x_k]_k \succ [x_k]_k$ for every $\lambda \in (0,1)$ and $[y^- \mu y_j]_j \prec [y_j]_j$ for every $\mu \in (0,1)$. Let $x^* = x^+ \lambda x_k$ and $y^* = y^- \mu y_j$ for some $\lambda, \mu \in (0,1)$. Denote

$$\mathbf{x}^*_{k,N} = (x_1, \ldots, x_{k-1}, x^*, x_{k+1}, \ldots, x_N, x_0, x_0, \ldots),$$

and

$$\mathbf{y}^*_{j,M} = (y_1, \ldots, y_{j-1}, y^*, y_{j+1}, \ldots, y_M, x_0, x_0, \ldots).$$

Then by Axiom I6, there exist T^-, T^+ such that

$$\mathbf{x}^*_{k,N} \succcurlyeq \mathbf{x} \succcurlyeq \mathbf{y} \succcurlyeq \mathbf{y}^*_{j,M}$$

for all $N \geq T^+$ and for all $M \geq T^-$. Since $\mathbf{x}^*_{k,N} \succcurlyeq \mathbf{y}^*_{j,M}$ and U represents $\succcurlyeq$ on X^* we have:

$$U\left(\mathbf{x}^*_{k,N}\right) \geq U\left(\mathbf{y}^*_{j,M}\right).$$

By Step 3 we know that $U(x_1, \ldots, x_{k-1}, x^*, x_{k+1} \ldots)$ and $U(y_1, \ldots, y_{j-1}, y^*, y_{j+1}, \ldots)$ converge, so

$$U(x_1, \ldots, x_{k-1}, x^*, x_{k+1}, \ldots) \geq U(y_1, \ldots, y_{j-1}, y^*, y_{j+1}, \ldots).$$

Recall that $x^* = x^+ \lambda x_k$ and $y^* = y^- \mu y_j$ for some $\lambda \in (0,1)$ and some $\mu \in (0,1)$. Since λ and μ are arbitrary, it follows that $U(\mathbf{x}) \geq U(\mathbf{y})$.

If it is not possible to find x^+, y^- such that $[x^+]_k \succ [x_k]_k$ and $[y^-]_j \prec [y_j]_j$, then either $w_t = 0$ for all t, in which case $U(\mathbf{x}) = U(\mathbf{y})$; or $x_t \succcurlyeq z'$ for all $z' \in X$ and all t with $w_t > 0$, in which case $U(\mathbf{x}) \geq U(\mathbf{y})$; or $z' \succcurlyeq y_t$ for all $z' \in X$ and all t with $w_t > 0$ in which case $U(\mathbf{x}) \geq U(\mathbf{y})$.

It is worth noting that as $\mathbf{x} \succcurlyeq \mathbf{y}$ implies $U(\mathbf{x}) \geq U(\mathbf{y})$, then, by Axiom I2 it follows that $w_t > 0$ for at least one t. Therefore, $\sum_{t=1}^{\infty} w_t > 0$. Normalizing by $1/\sum_{t=1}^{\infty} w_t$, we can assume that $\sum_{t=1}^{\infty} w_t = 1$.

Next, assume that $U(\mathbf{x}) \geq U(\mathbf{y})$. Suppose that it is possible to find k and $x^+, x^- \in$

X such that $\mathbf{x}^+_{k,N} \succcurlyeq \mathbf{x} \succcurlyeq \mathbf{x}^-_{k,N}$ for some fixed N. By mixture continuity, the set $\{\alpha \mid \mathbf{x}^+_{k,N}\alpha\mathbf{x}^-_{k,N} \succcurlyeq \mathbf{x}\}$ is closed. By assumption $\mathbf{x}^+_{k,N} \succcurlyeq \mathbf{x}$, so it follows that $\alpha = 1$ is included into the set. Analogously, the set $\{\beta \mid \mathbf{x} \succcurlyeq \mathbf{x}^+_{k,N}\beta\mathbf{x}^-_{k,N}\}$ is closed. In fact, $\beta = 0$ belongs to the set, as $\mathbf{x} \succcurlyeq \mathbf{x}^-_{k,N}$. Therefore, as both sets are closed, nonempty and form the unit interval, their intersection is nonempty. Hence, there exists λ such that $\mathbf{x} \sim \mathbf{x}^+_{k,N}\lambda\mathbf{x}^-_{k,N}$. Note that $\mathbf{x}^+_{k,N}\lambda\mathbf{x}^-_{k,N} = (x_1, \ldots, x_{k-1}, x^+\lambda x^-, x_{k+1} \ldots, x_N, x_0, x_0, \ldots)$. Let $x^+\lambda x^- = x^*$. Define $\mathbf{x}^*_{k,N} = (x_1, \ldots, x_{k-1}, x^*, x_{k+1}, \ldots, x_N, x_0, x_0, \ldots)$. Therefore, if there exist periods k, j and outcomes $x^+, x^-, y^+, y^- \in X$ such that $\mathbf{x}^+_{k,N} \succcurlyeq \mathbf{x} \succcurlyeq \mathbf{x}^-_{k,N}$ and $\mathbf{y}^+_{j,M} \succcurlyeq \mathbf{y} \succcurlyeq \mathbf{y}^-_{j,M}$ for some N and some M, we can find $\lambda, \mu \in [0, 1]$ such that $\mathbf{x} \sim \mathbf{x}^*_{k,N}$ and

$$\mathbf{y} \sim \mathbf{y}^*_{j,M} = (y_1, \ldots, y_{j-1}, y^*, y_{j+1}, \ldots, y_M, x_0, x_0, \ldots),$$

where $y^* = y^+\mu y^-$. We have already shown that if $\mathbf{x} \succcurlyeq \mathbf{y}$ then $U(\mathbf{x}) \geq U(\mathbf{y})$. From $\mathbf{x} \sim \mathbf{x}^*_{k,N}$ and $\mathbf{y} \sim \mathbf{y}^*_{j,M}$ it therefore follows that:

$$U\left(\mathbf{x}^*_{k,N}\right) = U(\mathbf{x}) \text{ and } U\left(\mathbf{y}^*_{j,M}\right) = U(\mathbf{y}).$$

Hence, from the assumption $U(\mathbf{x}) \geq U(\mathbf{y})$ we obtain:

$$U\left(\mathbf{x}^*_{k,N}\right) \geq U\left(\mathbf{y}^*_{j,M}\right).$$

Recall that U is an order-preserving function on X^*. Thus, $\mathbf{x}^*_{k,N} \succcurlyeq \mathbf{y}^*_{j,M}$. Since $\mathbf{x} \sim \mathbf{x}^*_{k,N}$ and $\mathbf{y} \sim \mathbf{y}^*_{j,M}$, we obtain $\mathbf{x} \succcurlyeq \mathbf{y}$.

Suppose now that there is no such k, j or outcomes x^+, x^-, y^+, y^- such that $\mathbf{x}^+_{k,N} \succcurlyeq \mathbf{x} \succcurlyeq \mathbf{x}^-_{k,N}$ and $\mathbf{y}^+_{j,M} \succcurlyeq \mathbf{y} \succcurlyeq \mathbf{y}^-_{j,M}$ for some N and some M. Then, using Axiom I6, we can conclude that either $x_t \in X^e$ for every t with $w_t > 0$ or $y_t \in X^e$ for every t with $w_t > 0$. Assume that there is only an upper bound to preferences; i.e.,

$$X^e \equiv \{z \in X \mid z \succcurlyeq z' \text{ for every } z' \in X\}.$$

Then $U(\mathbf{x}) \geq U(\mathbf{y})$ means that $x_t \in X^e$ whenever $w_t > 0$. Therefore, $U(\mathbf{x}) = U(\overline{\mathbf{x}})$, where $\overline{\mathbf{x}} = (\overline{x}, \overline{x}, \ldots)$ and $\overline{x} \in X^e$. Hence, it follows by monotonicity that $\mathbf{x} \succcurlyeq \mathbf{y}$. In the case when there is only a lower bound, i.e., $\underline{x} \in X^e \equiv \{z \in X \mid z' \succcurlyeq z \text{ for every } z' \in X\}$, the argument is similar.

Next, suppose that X is preference bounded above and below, i.e., there exist $\underline{x}, \overline{x} \in X^e$ with $\overline{x} \succcurlyeq x \succcurlyeq \underline{x}$ for every $x \in X$. Assume that $U(\mathbf{x}) \geq U(\mathbf{y})$. We need to demonstrate that $\mathbf{x} \succcurlyeq \mathbf{y}$. By monotonicity and continuity there exist $\lambda, \mu \in [0, 1]$ such that $\mathbf{x} \sim \overline{\mathbf{x}}\lambda\underline{\mathbf{x}}$ and $\mathbf{y} \sim \overline{\mathbf{x}}\mu\underline{\mathbf{x}}$. Since by assumption $U(\mathbf{x}) \geq U(\mathbf{y})$ and U represents the preference order on constant streams, we have $U(\overline{\mathbf{x}}\lambda\underline{\mathbf{x}}) \geq U(\overline{\mathbf{x}}\mu\underline{\mathbf{x}})$. By rearranging this inequality $(\lambda-\mu)(U(\overline{\mathbf{x}})-U(\underline{\mathbf{x}}))$, and using $U(\overline{\mathbf{x}}) > U(\underline{\mathbf{x}})$ it follows that $\lambda \geq \mu$. Therefore, as $\mathbf{x} \sim \overline{\mathbf{x}}\lambda\underline{\mathbf{x}}$ and $\mathbf{y} \sim \overline{\mathbf{x}}\mu\underline{\mathbf{x}}$ and $\lambda \geq \mu$, we conclude that $\mathbf{x} \succcurlyeq \mathbf{y}$.

Thus $(\mathbf{w}, u)$ is an AA representation for $\succcurlyeq$.

Step 5. Uniqueness of w_t. Assume that $(\mathbf{w}', u')$ is another AA representation. Then, for any t we have $w_t > 0$ if and only if $w'_t > 0$. Consider the set of all constant programs $\{\mathbf{x} \in X^\infty \mid \mathbf{x} = (a, a, \ldots), \text{ where } a \in X\}$, which is a mixture set. Applying $(\mathbf{w}', u')$ and $(\mathbf{w}, u)$ to this set we conclude that $u(a) > u(b)$ if and only if $u'(a) > u'(b)$ for every $a, b \in X$. By Theorem 1 [31] it implies that $u = Au' + B$ for some $A > 0$ and some B. Hence,

$$\sum_{t=1}^{\infty} w_t u(x_t) \geq \sum_{t=1}^{\infty} w_t u(y_t) \text{ if and only if } \sum_{t=1}^{\infty} w'_t u(x_t) \geq \sum_{t=1}^{\infty} w'_t u(y_t).$$

For any t, s with $t \neq s$ and any $x', x'' \in X$, let $[x', x'']_{t,s}$ denote the stream with x' in the t^{th} position, x'' in the s^{th} position and x_0 elsewhere. Fix t, s with $w_t > 0$ and $w_s > 0$. Using non-triviality, choose some $x^+, x^- \in X$ such that $x^+ \succ x^-$. Define $\mathbf{x} = [x^+, x^+]_{t,s}$, $\mathbf{y} = [x^+, x^-]_{t,s}$, $\mathbf{z} = [x^-, x^-]_{t,s}$. From the AA representation it follows that $\mathbf{x} \succ \mathbf{y} \succ \mathbf{z}$. By continuity of the AA representation there exists $\lambda \in (0,1)$ such that $\mathbf{y} \sim \mathbf{x}\lambda\mathbf{z}$. Applying the AA representation to $\mathbf{y} \sim \mathbf{x}\lambda\mathbf{z}$ we obtain

$$w_t u(x^+) + w_s u(x^-) = \lambda(w_t + w_s)u(x^+) + (1-\lambda)(w_t + w_s)u(x^-).$$

It follows that $(1-\lambda)w_t = \lambda w_s$. Similarly, $(1-\lambda)w'_t = \lambda w'_s$. Therefore, $w_t/w_s = w'_t/w'_s$. As this is true for any t, s, we obtain that $\mathbf{w} = C\mathbf{w}'$ for some $C > 0$. The sufficiency of the uniqueness conditions follows by routine arguments. $\square$

2.3 Discounted utility: finite case ($n < \infty$)

2.3.1 Exponential discounting

Recall that a preference $\succcurlyeq$ on X^n is represented by an exponentially discounted utility function if there exists a non-constant function $u\colon X \to \mathbb{R}$ and a parameter $\delta \in (0,1)$ such that

$$U(\mathbf{x}) = \sum_{t=1}^{n} \delta^{t-1} u(x_t).$$

If u is mixture linear (and non-constant), then we say that the pair (u, δ) provides an exponentially discounted expected utility representation.

Based on Theorem 2.1 it is easy to obtain such a representation. To do so, an adjustment of non-triviality and two additional axioms - impatience and stationarity - are required.

Axiom F2′. (Essentiality of period 1). There exist some $a, b \in X$ and some $\mathbf{x} \in X^n$ such that $(a, x_2, \ldots, x_n) \succ (b, x_2, \ldots, x_n)$.

Axiom F6. (Impatience). For all $a, b \in X$ if $a \succ b$, then for all $\mathbf{x} \in X^n$

$$(a, b, x_3, \ldots, x_n) \succ (b, a, x_3, \ldots, x_n).$$

Axiom F7. (Stationarity). The preference $(a, x_2, \ldots, x_n) \succcurlyeq (a, y_2, \ldots, y_n)$ holds if and only if $(x_2, \ldots, x_n, a) \succcurlyeq (y_2, \ldots, y_n, a)$ for every $a \in X$ and every $\mathbf{x}, \mathbf{y} \in X^n$.

It is not hard to see that essentiality of each period t follows from the essentiality of period 1 and the stationarity axiom.

Now the following result can be stated:

Theorem 2.3 (Exponential discounting). *The preferences $\succcurlyeq$ on X^n satisfy axioms F1, F2′, F3-F7 if and only if there exists an exponentially discounted expected utility representation (u, δ) for $\succcurlyeq$ on X^n. Moreover, (u', δ') is another exponentially discounted expected utility representations for $\succcurlyeq$ on X^n if and only if there are some $A > 0$ and some B such that $u = Au' + B$ and $\delta = \delta'$.*

Proof. It is straightforward to show that the axioms are implied by the representation. Conversely, suppose the axioms hold. Note that non-triviality follows from essentiality of period 1 and monotonicity.

By Theorem 1 we therefore know that $\succcurlyeq$ has an AA representation $(u, \mathbf{w})$. Define $\succcurlyeq'$ on X^{n-1} as follows:

$$(x_1, \ldots, x_{n-1}) \succcurlyeq' (y_1, \ldots, y_{n-1}) \Leftrightarrow (x_0, x_1, \ldots, x_{n-1}) \succcurlyeq (x_0, y_1, \ldots, y_{n-1}).$$

Then $\succcurlyeq'$ is represented by:

$$U'(\mathbf{x}) = w_2 u(x_1) + \ldots + w_n u(x_{n-1}).$$

Next, define $\succcurlyeq''$ on X^{n-1} as follows:

$$(x_1, \ldots, x_{n-1}) \succcurlyeq'' (y_1, \ldots, y_{n-1}) \Leftrightarrow (x_1, \ldots, x_{n-1}, x_0) \succcurlyeq (y_1, \ldots, y_{n-1}, x_0).$$

Then $\succcurlyeq''$ is represented by:

$$U''(\mathbf{x}) = w_1 u(x_1) + \ldots + w_{n-1} u(x_{n-1}).$$

According to stationarity, these preferences are equivalent ($\succcurlyeq' \equiv \succcurlyeq''$) with two different AA representations (U' and U''). Preference orders $\succcurlyeq' \equiv \succcurlyeq''$ satisfy the AA axioms on X^{n-1}. Recall that w_t are unique up to a scale. Hence, $w_{t+1} = \delta w_t$ for some $\delta > 0$ and it follows that

$$w_n = \delta w_{n-1} = \delta^2 w_{n-2} = \ldots = \delta^{n-t} w_t = \ldots = \delta^{n-1} w_1.$$

Since all periods are essential it is without loss of generality to set $w_1 = 1$. Then we obtain the following representation for $\succcurlyeq$ on X^n:

$$U(\mathbf{x}) = \sum_{t=1}^{n} \delta^{t-1} u(x_t), \quad \text{where } \delta > 0.$$

Since impatience holds: if $a \succ b$, then

$$(a, b, x_3, \ldots, x_n) \succ (b, a, x_3, \ldots, x_n).$$

From the representation it follows that:

$$u(a) + \delta u(b) > u(b) + \delta u(a),$$

or, equivalently,

$$(1 - \delta)\,(u(a) - u(b)) > 0.$$

As $u(a) > u(b)$, it is possible to conclude that $\delta \in (0, 1)$.

We now prove the uniqueness part of the theorem. Suppose that (u, δ) and (u', δ') both provide exponentially discounted expected utility representations for $\succcurlyeq$ on X^n. We need to show that $u = Au' + B$ for some $A > 0$ and $\delta = \delta'$. Indeed, since (u, δ) and (u', δ') both provide AA representations for $\succcurlyeq$ it follows that $u = Au' + B$ for some $A > 0$ and some B, and there is some $C > 0$ such that $\delta^{t-1} = C(\delta')^{t-1}$ for all t. Taking $t = 1$ we obtain $C = 1$, and hence $\delta = \delta'$. The sufficiency of the uniqueness conditions follows by routine arguments. $\square$

2.3.2 Semi-hyperbolic discounting

A preference $\succcurlyeq$ on X^n has an $SH(T)$ discounted utility representation if there exists a non-constant function $u\colon X \to \mathbb{R}$ and parameters $\beta_1 \le \beta_2 \le \ldots \le \beta_{T-1}$, and $\beta_t \in (0, 1]$

for all $t \leq T-1$ and $\delta \in (0,1)$ such that the following function represents $\succcurlyeq$:

$$U(\mathbf{x}) = u(x_1) + \beta_1 \delta u(x_2) + \beta_1\beta_2\delta^2 u(x_3) + \ldots + \beta_1\beta_2 \cdots \beta_{T-2}\delta^{T-2}u(x_{T-1})$$
$$+ \beta_1\beta_2 \cdots \beta_{T-1} \sum_{t=T}^{n} \delta^{t-1}u(x_t).$$

If u is mixture linear (and non-constant), then the function U is called an $SH(T)$ *discounted expected utility representation*. In this case, we say that $(u, \boldsymbol{\beta}, \delta)$ provides an $SH(T)$ discounted expected utility representation, where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_{T-1})$.

To obtain this form of discounting a number of modifications to the set of axioms is required. A stronger essentiality condition should be used:

Axiom F2″. (Essentiality of periods $1, \ldots, T$). There exist some $a, b \in X$ and some $\mathbf{x} \in X^n$ such that for every $t = 1, \ldots, T$:

$$(x_1, x_2, \ldots, x_{t-1}, a, x_{t+1}, \ldots, x_n) \succ (x_1, x_2, \ldots, x_{t-1}, b, x_{t+1}, \ldots, x_n).$$

The impatience axiom, which is used to guarantee $\delta \in (0,1)$, should be restated for the periods T and $T+1$:

Axiom F6′. (Impatience). For every $a, b \in X$ if $a \succ b$, then for every $\mathbf{x} \in X^n$:

$$(x_1, \ldots, x_{T-1}, a, b, x_{T+2}, \ldots, x_n) \succ (x_1, \ldots, x_{T-1}, b, a, x_{T+2}, \ldots, x_n).$$

The generalization requires relaxing the axiom of stationarity to stationarity from period T.

Axiom F7′. (Stationarity from period T). The preference

$$(x_1, \ldots, x_{T-1}, a, x_{T+1}, \ldots, x_n) \succcurlyeq (x_1, \ldots, x_{T-1}, a, y_{T+1}, \ldots, y_n)$$

holds if and only if

$$(x_1, \ldots, x_{T-1}, x_{T+1}, \ldots, x_n, a) \succcurlyeq (x_1, \ldots, x_{T-1}, y_{T+1}, \ldots, y_n, a)$$

for every $a \in X$ and every $\mathbf{x} \in X^n$.

Our final axiom is motivated by the notion of present bias. The axiom of present bias ([58, Axiom 10]) is usually stated for preference orders on X^∞ and involves trade-offs between two periods $\{1, 2\}$.

Present Bias ([58, Axiom 10])

For every $a, b, c, d, e \in X$ such that $a \succ c, b \prec d$, for all $\mathbf{x} \in X^\infty$:

$$\text{if } (e, a, b, e, \ldots) \sim (e, c, d, e, \ldots), \text{ then } (a, b, e, \ldots) \succcurlyeq (c, d, e, \ldots).$$

The present bias axiom can be informally described as follows. Suppose there are two equivalent consumption streams one of which has larger consumption at $t = 2$ but smaller consumption at $t = 3$ than the other, with consumptions at other periods being equal. Then if the consumption at period $t = 1$ is removed from both streams and both streams are shifted forward by one period, a decision-maker will prefer the stream with the bigger consumption at $t = 1$ but smaller consumption at $t = 2$, thus valuing present consumption ($t = 1$) more highly. In our framework, this axiom can be adapted to the finite case and extended so that present bias may arise between any periods $\{t, t+1\}$, where $t \leq T$. Suppose that there are two identical consumption streams that differ only in values at periods $\{t, t+1\}$, where $t \leq T$. Early bias between $\{t, t+1\}$ means that if the first stream has a bigger level of consumption at period t but smaller level of consumption at period $t+1$ than the second stream, then shifting the consumption at period $t-1$ to the last period and shifting all consumption from period t onwards forward by one period changes the preference in favour of the first consumption stream.

Axiom F8. (Early bias) For every $a, b, c, d \in X$ such that $a \succ c, b \prec d$, for all $\mathbf{x} \in X^n$ and every $t \leq T$ if

$$(x_1, \ldots, x_{t-1}, a, b, x_{t+2}, \ldots, x_n) \sim (x_1, \ldots, x_{t-1}, c, d, x_{t+2}, \ldots, x_n), \text{ then}$$

$$(x_1, \ldots, x_{t-2}, a, b, x_{t+2}, \ldots, x_n, x_{t-1}) \succcurlyeq (x_1, \ldots, x_{t-2}, c, d, x_{t+2}, \ldots, x_n, x_{t-1}).$$

The early bias axiom is also referred to as the extended present bias axiom.

Theorem 2.4 (Semi-hyperbolic discounting)**.** *The preferences $\succcurlyeq$ on X^n satisfy axioms F1, F2′, F3, F4, F5, F6′, F7′, F8 if and only if there exists an $SH(T)$ discounted expected utility representation $(u, \boldsymbol{\beta}, \delta)$ for $\succcurlyeq$ on X^n. Moreover, $(u', \boldsymbol{\beta'}, \delta')$ is another $SH(T)$ discounted expected utility representation for $\succcurlyeq$ on X^n if and only if there are some $A > 0$ and some B such that $u = Au' + B$ and $\delta = \delta'$, $\boldsymbol{\beta} = \boldsymbol{\beta'}$.*

Proof. It can be easily seen that the axioms are implied by the representation. Suppose that the axioms hold. As for Theorem 2.3, the conditions of AA representation are satisfied, so it follows that $\succcurlyeq$ has an AA representation $(\mathbf{w}, u)$. Define $\succcurlyeq'$ on X^{n-T} as follows:

$$(x_1, \ldots, x_{n-T}) \succcurlyeq' (y_1, \ldots, y_{n-T}) \Leftrightarrow$$

$$(x_0, \ldots, x_0, x_1, \ldots, x_{n-T}) \succcurlyeq (x_0, \ldots, x_0, y_1, \ldots, y_{n-T}).$$

Then $\succcurlyeq'$ is represented by:

$$U'(\mathbf{x}) = w_{T+1}u(x_1) + \ldots + w_n u(x_{n-T}).$$

Next, define $\succcurlyeq''$ on X^{n-T} as follows:

$$(x_1, \ldots, x_{n-T}) \succcurlyeq'' (y_1, \ldots, y_{n-T}) \Leftrightarrow$$

$$(x_0, \ldots, x_0, x_1, \ldots, x_{n-T}, x_0) \succcurlyeq (x_0, \ldots, x_0, y_1, \ldots, y_{n-T}, x_0).$$

Then $\succcurlyeq''$ is represented by:

$$U''(\mathbf{x}) = w_T u(x_1) + \ldots + w_{n-1}u(x_{n-T}).$$

According to stationarity from period T, the preferences are equivalent ($\succcurlyeq' \equiv \succcurlyeq''$) with two different AA representations (U' and U'').

Preference orders $\succcurlyeq' \equiv \succcurlyeq''$ satisfy the AA axioms on X^{n-T}. Recall that w_t are unique up to a scale. Hence, as essentiality holds for all t (which follows from Axiom F2$'$ and Axiom F7$'$), we have $w_{t+1} = \delta w_t$ for some $\delta > 0$ and hence

$$w_n = \delta w_{n-1} = \delta^2 w_{n-2} = \ldots = \delta^{n-t} w_t = \ldots = \delta^{n-T} w_T.$$

Therefore, $w_t = \delta^{t-T} w_T$ for all $t \geq T+1$. We, therefore, obtain the following representation for $\succcurlyeq$:

$$U(\mathbf{x}) = w_1 u(x_1) + \ldots + w_{T-1}u(x_{T-1}) + w_T \sum_{t=T}^{n} \delta^{t-T} u(x_t).$$

Because of the essentiality of the first period and uniqueness of u up to affine transformations, the function

$$\hat{U}(\mathbf{x}) = u(x_1) + \frac{w_2}{w_1}u(x_2) + \ldots + \frac{w_{T-1}}{w_1}u(x_{T-1}) + \frac{w_T}{w_1}\sum_{t=T}^{n} \delta^{t-T} u(x_t),$$

provides an alternative representation for $\succcurlyeq$ which will be used instead of $U(\mathbf{x})$ further in the proof.

Note that

$$\begin{aligned}
&\frac{w_3}{w_1} = \frac{w_3}{w_2} \cdot \frac{w_2}{w_1}, \\
&\ldots, \\
&\frac{w_T}{w_1} = \frac{w_T}{w_{T-1}} \cdot \frac{w_{T-1}}{w_{T-2}} \cdot \ldots \cdot \frac{w_2}{w_1}.
\end{aligned}$$

Let $\gamma_{t-1} = \frac{w_t}{w_{t-1}}$ for all $t \leq T$. Therefore,

$$\frac{w_2}{w_1} = \gamma_1,$$
$$\frac{w_3}{w_1} = \gamma_1\gamma_2,$$
$$\dots,$$
$$\frac{w_T}{w_1} = \gamma_1\gamma_2\dots\gamma_{T-1}.$$

With this notation:

$$\hat{U}(\mathbf{x}) = u(x_1) + \gamma_1 u(x_2) + \ldots + \gamma_1 \cdots \gamma_{T-2} u(x_{T-1}) + \gamma_1 \cdots \gamma_{T-1} \sum_{t=T}^{n} \delta^{t-T} u(x_t).$$

It is necessary to show that $\gamma_{t-1} = \beta_{t-1}\delta$ with $\beta_{t-1} \in (0,1]$ for all $t \leq T$.

Suppose that $t = T$. Choose $a, b, c, d \in X$ such that $u(b) < u(d)$, $u(a) > u(c)$ and

$$\gamma_1 \cdots \gamma_{T-1} u(a) + \gamma_1 \cdots \gamma_{T-1} \delta u(b) = \gamma_1 \cdots \gamma_{T-1} u(c) + \gamma_1 \cdots \gamma_{T-1} \delta u(d). \tag{2.3}$$

Since essentiality is satisfied for each period we can rearrange the equation (2.3):

$$\delta = \frac{u(a) - u(c)}{u(d) - u(b)}. \tag{2.4}$$

From (2.3) it also follows that

$$(x_1, \ldots, x_{T-1}, a, b, x_{T+2}, \ldots, x_n) \sim (x_1, \ldots, x_{T-1}, c, d, x_{T+2}, \ldots, x_n),$$

Therefore, by the early bias axiom:

$$(x_1, \ldots, x_{T-2}, a, b, x_{T+2}, \ldots, x_n, x_{T-1}) \succcurlyeq (x_1, \ldots, x_{T-2}, c, d, x_{T+2}, \ldots, x_n, x_{T-1}).$$

Thus we obtain:

$$\gamma_1 \cdots \gamma_{T-2} u(a) + \gamma_1 \cdots \gamma_{T-1} u(b) \geq \gamma_1 \cdots \gamma_{T-2} u(c) + \gamma_1 \cdots \gamma_{T-1} u(d).$$

Since the essentiality condition is satisfied for each period we can rearrange this inequality:

$$\gamma_{T-1} \leq \frac{u(a) - u(c)}{u(d) - u(b)}. \tag{2.5}$$

Comparing (2.4) to (2.5) we conclude that $\delta \geq \gamma_{T-1}$, therefore, $\gamma_{T-1} = \beta_{T-1}\delta$, where

$\beta_{T-1} \in (0, 1]$.

Analogously, suppose that $t = T-1$. Choose $a', b', c', d' \in X$ such that $u(b') < u(d')$, $u(a') > u(c')$ and

$$\gamma_1 \cdots \gamma_{T-2} u(a') + \gamma_1 \cdots \gamma_{T-1} u(b') = \gamma_1 \cdots \gamma_{T-2} u(c') + \gamma_1 \cdots \gamma_{T-1} u(d'), \tag{2.6}$$

where the last equality can be rewritten as follows (since essentiality is satisfied):

$$\gamma_{T-1} = \frac{u(a') - u(c')}{u(d') - u(b')}. \tag{2.7}$$

Then (2.6), early bias and essentiality of each period imply that

$$\gamma_{T-2} \leq \frac{u(a') - u(c')}{u(d') - u(b')}. \tag{2.8}$$

It follows from (2.7) and (2.8) that $\gamma_{T-2} \leq \gamma_{T-1}$. Therefore, $\gamma_{T-2} = \beta'_{T-2}\gamma_{T-1}$, where $\beta'_{T-2} \in (0, 1]$. Recall that $\gamma_{T-1} = \beta_{T-1}\delta$. Hence,

$$\gamma_{T-2} = \beta'_{T-2}\beta_{T-1}\delta = \beta_{T-2}\delta,$$

where $\beta_{T-2} = \beta'_{T-2}\beta_{T-1}$ and $\beta_{T-2} \in (0, 1]$ as both $\beta'_{T-2} \in (0, 1]$ and $\beta_{T-1} \in (0, 1]$. Note also that $\beta_{T-2} \leq \beta_{T-1}$.

Using the early bias axiom repeatedly for $t < T - 1$ we obtain $\gamma_{t-1} = \beta_{t-1}\delta$ with $\beta_{t-1} \in (0, 1]$ for all $t \leq T$ and $\beta_1 \leq \beta_2 \leq \ldots \leq \beta_{T-1}$. Hence,

$$\begin{aligned} \hat{U}(\mathbf{x}) &= u(x_1) + \beta_1\delta u(x_2) + \beta_1\beta_2\delta^2 u(x_3) + \ldots + \beta_1\beta_2 \cdots \beta_{T-2}\delta^{T-2} u(x_{T-1}) \\ &\quad + \beta_1\beta_2 \cdots \beta_{T-1} \sum_{t=T}^{n} \delta^{t-1} u(x_t). \end{aligned}$$

To show that $\delta \in (0, 1)$ the impatience axiom should be applied. For every $a, b \in X$ if $a \succ b$, then for every $\mathbf{x} \in X^n$

$$(x_1, \ldots, x_{T-1}, a, b, x_{T+2}, \ldots, x_n) \succ (x_1, \ldots, x_{T-1}, b, a, x_{T+2}, \ldots, x_n).$$

Then

$$\beta_1 \cdots \beta_{T-1}\delta^{T-1} u(a) + \beta_1 \cdots \beta_{T-1}\delta^T u(b) > \beta_1 \cdots \beta_{T-1}\delta^{T-1} u(b) + \beta_1 \cdots \beta_{T-1}\delta^T u(a).$$

Therefore, due to essentiality of each period:

$$(1 - \delta)(u(a) - u(b)) > 0.$$

Hence, $\delta \in (0,1)$.

We now prove uniqueness. Suppose that $(u, \boldsymbol{\beta}, \delta)$ and $(u', \boldsymbol{\beta'}, \delta')$ both provide $SH(T)$ discounted expected utility representations for $\succcurlyeq$ on X^n. Let $D(t)$ and $D'(t)$ be semi-hyperbolic discount functions for given $\boldsymbol{\beta}, \delta$ and $\boldsymbol{\beta'}, \delta'$, respectively. Since $(u, \boldsymbol{\beta}, \delta)$ and $(u', \boldsymbol{\beta'}, \delta')$ both provide AA representations for $\succcurlyeq$ it follows that $u = Au' + B$ for some $A > 0$ and some B, and there is some $C > 0$ such that $D(t) = C \cdot D'(t)$ for all t. Taking $t = 1$ we obtain $C = 1$, and hence, letting $t = 2, 3, \ldots, T$ we get $\beta_t \delta = \beta'_t \delta'$ for all $t \leq T$. Finally, letting $t = T + 1$ we conclude that $\delta = \delta'$. Therefore, $\boldsymbol{\beta} = \boldsymbol{\beta'}$. The sufficiency of the uniqueness conditions follows by routine arguments. □

2.4 Discounted utility: infinite case ($n = \infty$)

2.4.1 Exponential discounting

Based on the AA representation for the preferences over infinite consumption streams (Theorem 2.2), with some strengthening of non-triviality (Axiom I2) and the addition of a suitable stationarity axiom, discounting functions in an exponential form can be obtained. The impatience axiom is not needed since convergence (Axiom I6) plays its role.

Axiom I2′. (Essentiality of period 1). There exist some $a, b \in X$ such that $[a]_1 \succ x_0 \succ [b]_1$.

Axiom I7. (Stationarity). The preference $(a, x_1, x_2, \ldots) \succcurlyeq (a, y_1, y_2, \ldots)$ holds if and only if $(x_1, x_2, \ldots) \succcurlyeq (y_1, y_2, \ldots)$ for every $a \in X$ and every $\mathbf{x}, \mathbf{y} \in X^\infty$.

Theorem 2.5 (Exponential discounting)**.** *The preferences $\succcurlyeq$ on X^∞ satisfy axioms I1, I2′, I3-I7 if and only if there exists an exponentially discounted expected utility representation (u, δ) for $\succcurlyeq$ on X^∞. Moreover, (u', δ') is another exponentially discounted expected utility representation for $\succcurlyeq$ on X^∞ if and only if there are some $A > 0$, some B and some $C > 0$ such that $u = Au' + B$ and $\delta = \delta'$.*

Proof. The necessity of the axioms is straightforward. The proof of sufficiency follows the steps of the proof of Theorem 2.3 with $n = \infty$. Applying Theorem 2.2 to the preferences satisfying the stationarity axiom we obtain the representation:

$$U(\mathbf{x}) = \sum_{t=1}^{\infty} \delta^{t-1} u(x_t),$$

where $\delta > 0$ and $\mathbf{x} \in X^\infty$.

Next, instead of using the impatience axiom as it is done in the finite case, the convergence axiom is applied. Take a constant stream $\mathbf{a} = (a, a, \ldots)$, such that $u(a) \neq 0$. Then,

$$U(\mathbf{a}) = \sum_{t=1}^{\infty} \delta^{t-1} u(a) = u(a) \sum_{t=1}^{\infty} \delta^{t-1},$$

Convergence requires $\delta < 1$. The proof of the uniqueness claims is analogous to Theorem 2.3. □

2.4.2 Semi-hyperbolic discounting

The extension of semi-hyperbolic discounting to the case where $n = \infty$ is easily obtained.

Axiom I2″. (Essentiality of periods $1, \ldots, T$). For some $a, b \in X$ we have $[a]_t \succ x_0 \succ [b]_t$ for every $t = 1, \ldots, T$.

The generalization requires relaxing the axiom of stationarity to stationarity from period T.

Axiom I7′. (Stationarity from period T). The preference

$$(x_1, \ldots, x_{T-1}, a, x_{T+1}, \ldots) \succcurlyeq (x_1, \ldots, x_{T-1}, a, y_{T+1}, \ldots)$$

holds if and only if

$$(x_1, \ldots, x_{T-1}, x_{T+1}, \ldots) \succcurlyeq (x_1, \ldots, x_{T-1}, y_{T+1}, \ldots)$$

for every $a \in X$, and every $\mathbf{x} \in X^{\infty}$.

As in the finite case the addition of the early bias axiom is needed. Consider two consumption streams that differ only in values at periods $\{t, t+1\}$, where $t \leq T$. Early bias between $\{t, t+1\}$ means that if the first stream has more consumption at period t but less consumption at period $t+1$ than the second stream, then dropping the consumption at $t-1$ from both streams and advancing consumption from period t onwards by one period results in the first consumption stream being preferred to the second consumption stream.

Axiom I8. (Early bias) For every $a, b, c, d \in X$ such that $a \succ c, b \prec d$, and for all $\mathbf{x} \in X^{\infty}$ and every $t \leq T$

$$\text{if } (x_1, \ldots, x_{t-1}, a, b, x_{t+2}, \ldots) \sim (x_1, \ldots, x_{t-1}, c, d, x_{t+2}, \ldots), \text{ then}$$

$$(x_1, \ldots, x_{t-2}, a, b, x_{t+2}, \ldots) \succcurlyeq (x_1, \ldots, x_{t-2}, c, d, x_{t+2}, \ldots).$$

Theorem 2.6 (Semi-hyperbolic discounting)**.** *The preferences $\succcurlyeq$ on X^∞ satisfy axioms I1, I2′, I3-I6, I7′, I8 if and only if there exists an $SH(T)$ discounted expected utility representation $(u, \boldsymbol{\beta}, \delta)$ for $\succcurlyeq$ on X^∞. Moreover, $(u', \boldsymbol{\beta}', \delta')$ is another $SH(T)$ discounted expected utility representation for $\succcurlyeq$ on X^∞ if and only if there are some $A > 0$ and some B such that $u = Au' + B$ and $\delta = \delta'$, $\boldsymbol{\beta} = \boldsymbol{\beta}'$.*

Proof. The necessity of the axioms is obviously implied by the representation. The proof of sufficiency is analogous to the finite case. Applying Theorem 2.2 and stationarity from period T we get the representation:

$$U(\mathbf{x}) = w_1 u(x_1) + \ldots + w_{T-1} u(x_{T-1}) + w_T \sum_{t=T}^{\infty} \delta^{t-T} u(x_t).$$

Next, dividing by $w_1 > 0$ and introducing the notation $\frac{w_t}{w_{t-1}} = \gamma_{t-1} > 0$, where $t \le T$, the representation becomes

$$\hat{U}(\mathbf{x}) = u(x_1) + \gamma_1 u(x_2) + \ldots + \gamma_1 \cdots \gamma_{T-2} u(x_{T-1}) + \gamma_1 \cdots \gamma_{T-1} \sum_{t=T}^{\infty} \delta^{t-T} u(x_t).$$

Using essentiality of each period and the early bias axiom repeatedly, we demonstrate that $\gamma_{t-1} = \beta_{t-1}\delta$ with $\beta_{t-1} \in (0, 1]$ for all $t \le T$ and $\beta_1 \le \beta_2 \le \ldots \le \beta_{T-1}$. Therefore,

$$\begin{aligned}\hat{U}(\mathbf{x}) = {} & u(x_1) + \beta_1 \delta u(x_2) + \beta_1 \beta_2 \delta^2 u(x_3) + \ldots + \beta_1 \beta_2 \cdots \beta_{T-2} \delta^{T-2} u(x_{T-1}) \\ & + \beta_1 \beta_2 \cdots \beta_{T-1} \sum_{t=T}^{\infty} \delta^{t-1} u(x_t).\end{aligned}$$

Finally, to show that $\delta \in (0, 1)$, take a constant stream $\mathbf{a} = (a, a, \ldots)$, such that $u(a) \neq 0$. Then,

$$\begin{aligned}\hat{U}(\mathbf{a}) &= u(a) + \beta_1 \delta u(a) + \ldots + \beta_1 \cdots \beta_{T-2} \delta^{T-2} u(a) + \beta_1 \cdots \beta_{T-1} \sum_{t=T}^{\infty} \delta^{t-1} u(a) \\ &= u(a) \left(1 + \beta_1 \delta + \ldots + \beta_1 \cdots \beta_{T-2} \delta^{T-2} + \beta_1 \cdots \beta_{T-1} \sum_{t=T}^{\infty} \delta^{t-1} \right).\end{aligned}$$

Convergence requires $\delta < 1$.

The proof of the uniqueness claims is analogous to Theorem 2.4. □

2.5 Discussion

A number of axiomatizations of exponential and quasi-hyperbolic discounting have been suggested by different authors. In fact, all the axiomatizations use different assumptions and there is no straightforward transformation from one model to another. In this chapter we provided an alternative approach to get a time separable discounted utility representation, showing that Anscombe and Aumann's result can be exploited as a common background for axiomatizing exponential and quasi-hyperbolic discounting in both finite and infinite time horizons. In addition, we demonstrated that the axiomatization of quasi-hyperbolic discounting can be easily extended to $SH(T)$.

A key distinguishing feature of our set-up is the mixture set structure for X and the use of the mixture independence condition. An essential question, however, is whether mixture independence is normatively compelling in a time preference context, because states are mutually exclusive whereas time periods are not. It is worth mentioning that the temporal interpretation of the AA framework was also used by Wakai [79] to axiomatize an entirely different class of preferences, which exhibit a desire to spread bad and good outcomes evenly over time.

Commonly, the condition of joint independence is used to establish additive separability in time-preference models. Given $A \subseteq T$, where $T = \{1, \ldots, n\}$, and $\mathbf{x}, \mathbf{y} \in X^n$, define $\mathbf{x}_A\mathbf{y}$ as follows: $(x_A y)_t$ is x_t if $t \in A$ and y_t otherwise. The preference order $\succcurlyeq$ satisfies *joint independence* if for every $A \subseteq T$ and for every $\mathbf{x}, \mathbf{x}', \mathbf{y}, \mathbf{y}' \in X^n$:

$$\mathbf{x}_A\mathbf{y} \succcurlyeq \mathbf{x}'_A\mathbf{y} \text{ if and only if } \mathbf{x}_A\mathbf{y}' \succcurlyeq \mathbf{x}'_A\mathbf{y}'.$$

Joint independence is used to obtain an additively separable representation by Debreu [23], so we will sometimes refer to it as a Debreu-type independence condition. It is known that mixture independence implies joint independence ([37]), but whether joint independence (with some other plausible conditions) implies mixture independence is yet to be determined.

In fact, we are not the first to use a mixture-type independence condition in the context of time preferences. Wakai [79] also does so, though he uses the weaker form of *constant independence* introduced by Gilboa and Schmeidler [35].

A version of the mixture independence condition can also be formulated in a Savage environment ([73]) without objective probabilities, as discussed in [38]. Olea and Strzalecki [58] use precisely this version of mixture independence in one of their axiomatizations of quasi-hyperbolic discounting. For every $x, y \in X$ let us write (x, y) for $(x, y, y, \ldots) \in X^\infty$. Let $m(x_1, y_1)$ denote some $c \in X$ satisfying $(x_1, y_1) \sim (c, c)$. For any streams (x_1, x_2) and (z_1, z_2) the consumption stream $(m(x_1, z_1), m(x_2, z_2))$ is called a subjective mixture of (x_1, x_2) and (z_1, z_2). Olea and Strzalecki's version of the mixture independence axiom (their Axiom I2) is as follows: for every $x_1, x_2, y_1, y_2, z_1, z_2 \in X$ if $(x_1, x_2) \succcurlyeq (y_1, y_2)$, then

$$(m(x_1, z_1), m(x_2, z_2)) \succcurlyeq (m(y_1, z_1), m(y_2, z_2))$$

and

$$(m(z_1, x_1), m(z_2, x_2)) \succcurlyeq (m(z_1, y_1), m(z_2, y_2)) \, .$$

In other words, if a consumption stream (x_1, x_2) is preferred to a stream (y_1, y_2), then subjectively mixing each stream with (z_1, z_2) does not affect the preference.

In their axiomatization of quasi-hyperbolic discounting Olea and Strzalecki invoke their mixture independence condition (Axiom 12) as well as Debreu-type independence conditions. The latter are used to obtain a representation in the form

$$\mathbf{x} \succcurlyeq \mathbf{y} \text{ if and only if } u(x_1) + \sum_{t=2}^{\infty} \delta^{t-1} v(x_t) \geq u(y_1) + \sum_{t=2}^{\infty} \delta^{t-1} v(y_t),$$

then their Axiom 12 is used to ensure $v = \beta u$.[2]

Hayashi [44] and Epstein [26] considered preferences over lotteries over consumption streams. In their framework X^∞ is the set of non-stochastic consumption streams, where X is required to be a compact connected separable metric space. Denote the set of probability measures on the Borel σ-algebra defined on X^∞ as $\Delta(X^\infty)$. It is useful to note that our setting is the restriction of the Hayashi [44] and Epstein [26] set-up to product measures, i.e., to $\Delta(X)^\infty \subset \Delta(X^\infty)$. The axiomatization systems of Hayashi [44] and Epstein [26] are based on the assumptions of expected utility theory. The existence of a continuous and bounded vNM utility index $U\colon \Delta(X^\infty) \to \mathbb{R}$ is stated as one of the axioms. A set of necessary and sufficient conditions for this is provided by Grandmont [36], and includes the usual vNM independence condition on $\Delta(X^\infty)$: for every $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \Delta(X^\infty)$ and any $\alpha \in [0, 1]$, $\mathbf{x} \sim \mathbf{y}$ implies $\alpha\mathbf{x} + (1-\alpha)\mathbf{z} \sim \alpha\mathbf{y} + (1-\alpha)\mathbf{z}$.

Obviously, this independence condition is not strong enough to deliver joint independence of time periods, which is why additional assumptions of separability are needed. Two further Debreu-type independence conditions are required for exponential discounting:

- independence of stochastic outcomes in periods $\{1, 2\}$ from deterministic outcomes in $\{3, 4, \ldots\}$,

[2] As pointed out above, mixture independence stated for n periods implies joint independence for n periods. Hence, this raises the obvious question of whether it is possible to use an n-period version of the subjective mixture independence axiom to obtain a time separable discounted utility representation without the need for the Debreu-type independence conditions.

- independence of stochastic outcomes in periods $\{2, 3, \ldots\}$ from deterministic outcomes in period $\{1\}$.

To obtain quasi-hyperbolic discounting two additional Debreu-type independence conditions should be satisfied:

- independence of stochastic outcomes in periods $\{2, 3\}$ from deterministic outcomes in periods $\{1\}$ and $\{4, \ldots\}$,
- independence of stochastic outcomes in periods $\{3, 4, \ldots\}$ from deterministic outcomes in periods $\{1, 2\}$.

It is easy to see that these axioms applied to the non-stochastic consumption streams are analogous to the Debreu-type independence conditions used in [17] and [58].

In summary, to get a discounted utility representation with the discount function in either exponential and quasi-hyperbolic form separability must be assumed. The mixture independence axiom appears to be a strong assumption, however, it gives the desired separability without the need for additional Debreu-type independence conditions.

3

Aggregating time preferences with decreasing impatience

As discussed in Chapter 1, time preferences can vary significantly among decision-makers. This becomes an important issue when a decision is to be made by a group of decision-makers. For this reason, it is important to understand the properties of aggregated, or average, discount functions. Another reason, also discussed in Chapter 1, is the possibility that there is uncertainty about the appropriate discount function that a decision-maker should apply. Existing results in the literature on preference aggregation were introduced in Section 1.1.5. This chapter is focused on aggregating time preferences that exhibit decreasing impatience, as decreasing impatience remains the most common finding in experiments on intertemporal choice.

The chapter is organized as follows. In Section 3.1 we present the preliminaries on preferences and some essential results on (log-)convexity of functions. We proceed in Section 3.2 to a discussion of the comparative DI notion and define some instruments for such comparison. In Section 3.3 we analyse the behaviour of mixtures of discount functions. Section 3.4 offers a summary and some discussion of the results.

3.1 Preliminaries

In this section we introduce the framework for our investigation and define the two key concepts used in this chapter: strong present bias and strictly decreasing impatience of preferences. We prove that these two concepts coincide when the Fishburn-Rubinstein axioms [32] for a discounted utility representation are satisfied. Taking our lead from Pratt [61] and Arrow [9], these concepts are discussed in terms of log-convexity of discount functions. Necessary results and definitions concerning (log-)convex functions are given in the next section.

3.1.1 Convexity and log-convexity

Convexity and log-convexity play an important role in the theory of discounting. Let I be an interval (finite or infinite) of real numbers. A function $f\colon I \to \mathbb{R}$ is *convex* if for any two points $x, y \in I$ and any $\lambda \in [0,1]$ it holds that:

$$f\left(\lambda x + (1-\lambda)\, y\right) \leq \lambda f(x) + (1-\lambda) f(y).$$

A function f is *strictly convex* if

$$f(\lambda x + (1-\lambda) y) < \lambda f(x) + (1-\lambda) f(y)$$

for any $x, y \in I$ such that $x \neq y$ and any $\lambda \in (0,1)$. If f is twice differentiable convexity is equivalent to $f'' \geq 0$, and strict convexity is equivalent to two conditions: the function f'' is nonnegative on I and the set $\{x \in I \mid f''(x) = 0\}$ contains no non-trivial interval [76].

The following equivalent definition of a (strictly) convex function is well known. A function $f\colon I \to \mathbb{R}$ is (strictly) convex if for every $x, y, v, z \in I$ such that $x - y = v - z > 0$ and $y > z$ we have

$$f(x) - f(y) \leq [<] f(v) - f(z).$$

Convexity is preserved under composition of functions, as shown in the following lemma, whose straightforward proof is omitted:

Lemma 3.1. *Let $f_1\colon I \to \mathbb{R}$ be a non-decreasing and convex function and $f_2\colon I \to \mathbb{R}$ be a convex function, such that the range of f_2 is contained in the domain of f_1. Then the composition $f = f_1 \circ f_2$ is a convex function. If, in addition, f_1 is strictly increasing, and either f_1 or f_2 is strictly convex, then f is also strictly convex.*

A function $f\colon I \to \mathbb{R}$ is called *log-convex* if $f(x) > 0$ for all $x \in I$ and $\ln(f)$ is convex. It is called *strictly log-convex* if $\ln(f)$ is strictly convex. If follows that if f is a (strictly positive) twice differentiable function, then log-convexity of f is equivalent to the condition $f''f - (f')^2 \geq 0$, while strict log-convexity of f requires, in addition, that the set

$$\{\, x \in I \mid f''(x)f(x) - [f'(x)]^2 = 0 \,\}$$

contains no non-trivial interval. Log-convexity can also be expressed without using logarithms [19]. A function $f\colon I \to \mathbb{R}$ is log-convex if and only if $f(x) > 0$ for all $x \in I$ and for all $x, y \in I$ and $\lambda \in [0, 1]$ we have:

$$f(\lambda x + (1-\lambda)y) \leq f(x)^{\lambda} f(y)^{1-\lambda}. \tag{3.1}$$

The function f is strictly log-convex if inequality (3.1) is strict when $x \neq y$ and $\lambda \in (0, 1)$.

It is well known that the sum of two log-convex functions is log-convex [10, Theorem 1.8]. We have included a proof of the following variation on this result for completeness.

Lemma 3.2. *Let $f, g\colon I \to \mathbb{R}$ be functions with f strictly log-convex and g log-convex. Then the sum $f + g$ is strictly log-convex.*

Proof. Since $f(x) > 0$ and $g(x) > 0$ for all $x \in I$, we have $(f + g)(x) > 0$ for all $x \in I$. Let $x, y \in I$ such that $x \neq y$ and let $\lambda \in (0, 1)$. We must show that

$$f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) < (f(x) + g(x))^{\lambda}(f(y) + g(y))^{1-\lambda}.$$

Since f is strictly log-convex, we have

$$f(\lambda x + (1-\lambda)y) < f(x)^{\lambda} f(y)^{1-\lambda}. \tag{3.2}$$

Analogously, since $g(x)$ is log-convex:

$$g(\lambda x + (1-\lambda)y) \leq g(x)^{\lambda} g(y)^{1-\lambda}. \tag{3.3}$$

Summing (3.2) and (3.3) we obtain:

$$f(\lambda x + (1-\lambda)y) + g(\lambda x + (1-\lambda)y) < f(x)^{\lambda} f(y)^{1-\lambda} + g(x)^{\lambda} g(y)^{1-\lambda}.$$

Denote $a = f(x), b = f(y), c = g(x), d = g(y)$. Note that $a, b, c, d > 0$. To prove the claim of the lemma, it is sufficient to show that:

$$a^{\lambda} b^{1-\lambda} + c^{\lambda} d^{1-\lambda} \leq (a + c)^{\lambda}(b + d)^{1-\lambda}. \tag{3.4}$$

Since $(a+c)^{\lambda}(b+d)^{1-\lambda} > 0$ we can divide both parts of (3.4) by this expression to get

$$\left(\frac{a}{a+c}\right)^{\lambda}\left(\frac{b}{b+d}\right)^{1-\lambda} + \left(\frac{c}{a+c}\right)^{\lambda}\left(\frac{d}{b+d}\right)^{1-\lambda} \leq 1.$$

By the Weighted AM-GM inequality [22, Theorem 7.6, p. 74]:

$$\left(\frac{a}{a+c}\right)^{\lambda}\left(\frac{b}{b+d}\right)^{1-\lambda} \leq \lambda\frac{a}{a+c} + (1-\lambda)\frac{b}{b+d}$$

and

$$\left(\frac{c}{a+c}\right)^{\lambda}\left(\frac{d}{b+d}\right)^{1-\lambda} \leq \lambda\frac{c}{a+c} + (1-\lambda)\frac{d}{b+d}.$$

Hence,

$$\left(\frac{a}{a+c}\right)^{\lambda}\left(\frac{b}{b+d}\right)^{1-\lambda} + \left(\frac{c}{a+c}\right)^{\lambda}\left(\frac{d}{b+d}\right)^{1-\lambda} \leq \lambda + (1-\lambda) = 1,$$

which proves the statement in the lemma. □

One of the important definitions which will be frequently used throughout the chapter is that of a convex transformation. We say that f_1 is a *(strictly) convex transformation* of f_2 if there exists a (strictly) convex function f such that $f_1(x) = (f \circ f_2)(x) = f(f_2(x))$.

Lemma 3.3. *Let $f_1, f_2 \colon I \to \mathbb{R}$ such that f_2^{-1} exists. Then f_1 is a (strictly) convex transformation of f_2 if and only if the composition $f_1 \circ f_2^{-1}$ is (strictly) convex.*

Proof. See [61]. □

Recall also that a function $f \colon I \to \mathbb{R}$ is called *concave* if and only if $-f$ is convex. Thus a function $f \colon I \to \mathbb{R}$ is *log-concave* if and only if $1/f$ is log-convex. Therefore, the definitions and results stated in this section can be easily adapted for (log-)concavity.

3.1.2 Preferences

Let X be the set of outcomes. In this chapter we will assume that X is an interval of non-negative real numbers containing 0. The natural interpretation is that outcomes are monetary (for an infinitely divisible currency) but this is not essential. Let $T = [0, \infty)$ be a set of points in time, where 0 corresponds to the present moment. The set $\mathcal{A}_1 = X \times T$ will be identified with the set of dated outcomes, i.e., a pair $(x, t) \in X \times T$ is understood as a dated outcome, in which a decision-maker receives x at time t and the "neutral" outcome, 0, at all other time periods in $T \setminus \{t\}$.

Suppose that a decision-maker has a preference order $\succcurlyeq$ on the set of dated outcomes with $\succ$ expressing strict preference and $\sim$ indifference. We say that a utility function $U\colon X \times T \to \mathbb{R}$ *represents* the preference order $\succcurlyeq$, if for all $x, y \in X$ and all $t, s \in T$ we have $(x,t) \succcurlyeq (y,s)$ if and only if $U(x,t) \geq U(y,s)$. This is a *discounted utility (DU) representation* if

$$U(x,t) = D(t)u(x), \tag{3.5}$$

where $u\colon X \to \mathbb{R}$ is a continuous and strictly increasing function with $u(0) = 0$, and $D\colon T \to (0,1]$ is continuous and strictly decreasing such that $D(0) = 1$ and $\lim_{t\to\infty} D(t) = 0$.

The function u is called the *instantaneous utility function*, and D is called the *discount function* associated with $\succcurlyeq$. We say that the pair (u, D) provides a *DU representation* for $\succcurlyeq$. Fishburn and Rubinstein [32] provide an axiomatic foundation for a DU representation.[1]

We assume that $\succcurlyeq$ *has a discounted utility representation* throughout the chapter.

As D is strictly decreasing, our decision-maker is always impatient. However, as time goes by, her impatience may increase or decrease.

Definition 3.4 ([63]). *The preference order $\succcurlyeq$ exhibits (strictly) decreasing impatience (DI) if for all $\sigma > 0$, all $0 \leq t < s$ and all outcomes $y > x > 0$, the equivalence $(x,t) \sim (y,s)$ implies $(x, t+\sigma) \preccurlyeq [\prec]\ (y, s+\sigma)$.*

Increasing impatience (II) can be defined by reversing the final preference ranking in Definition 3.4. However, we focus on DI preferences in the present chapter, since this appears to be the empirically relevant case [33]. As in the previous sentence, we also use the acronym "DI" interchangeably as a noun ("decreasing impatience") and an adjective ("decreasingly impatient"), relying on context to indicate the intended meaning.

In case the preference order $\succcurlyeq$ has a DU representation, the characterisation of DI in terms of the discount function is well-known.[2]

Proposition 3.5 ([43, 63]). *Let $\succcurlyeq$ be a preference order having DU representation with the discount function D. The following conditions are equivalent:*

- *The preference order $\succcurlyeq$ exhibits (strictly) DI;*
- *D is (strictly) log-convex on $[0, \infty)$.*

[1] A list of their axioms is given in Chapter 1.

[2] The proof in [43, Theorem 3.3] can be easily adapted to demonstrate an analogous result for increasing impatience: the preference order $\succcurlyeq$ exhibits (strictly) II if and only if D is (strictly) log-concave on $[0, \infty)$.

We say that discount function D is (strictly) DI if there exists a preference order $\succcurlyeq$ that exhibits (strictly) DI and has a DU representation with discount function D.

We next show that a preference order $\succcurlyeq$ with a DU representation exhibits strictly DI if and only if it exhibits strong present bias in the following sense:

Definition 3.6 (Strong Present Bias). *The preference order exhibits strong present-bias if*

(i) $(x, t) \preccurlyeq (y, s)$ *implies* $(x, t+\sigma) \preccurlyeq (y, s+\sigma)$ *for every* x, y, *every* $\sigma > 0$ *and every* $s, t \in T$ *such that* $s > t \geq 0$; *and*

(ii) for every $s, t \in T$ *with* $s > t \geq 0$ *and every* $\sigma > 0$ *there exist* x^* *and* y^* *such that* $(x^*, t+\sigma) \prec (y^*, s+\sigma)$ *and* $(x^*, t) \succ (y^*, s)$.

Definition 3.6 is equivalent to the concept of "present bias" introduced by Jackson and Yariv in [46] but adapted to continuous time.

Proposition 3.7 gives conditions which are equivalent to strong present bias for preferences with a DU representation:

Proposition 3.7. *Suppose that* $\succcurlyeq$ *has a DU representation. Then the first condition of Definition 3.6 is equivalent to convexity of* $\ln D(t)$*; while the second condition of Definition 3.6 is equivalent to strict convexity of* $\ln D(t)$.

Proof. We start by proving the first equivalence. Since a DU representation exists, the first condition is equivalent to:

$$u(x)D(t) \leq u(y)D(s) \quad \text{implies} \quad u(x)D(t+\sigma) \leq u(y)D(s+\sigma)$$

for every x, y, every $\sigma > 0$ and every $s, t \in T$ with $s > t \geq 0$. This may be rewritten as follows:

$$u(x) \leq \frac{D(s)}{D(t)}u(y) \quad \text{implies} \quad u(x) \leq \frac{D(s+\sigma)}{D(t+\sigma)}u(y).$$

Since $(y, s) \succcurlyeq (x, t)$, $s > t$ and D is strictly decreasing it follows that $u(y) > u(x)$. As $u(0) = 0$ and u is strictly increasing we deduce that $u(y) > 0$. Since u is continuous, x and y can be chosen so that $u(x)/u(y)$ takes any value in $[0, 1)$. We therefore have:

$$\frac{D(s)}{D(t)} \leq \frac{D(s+\sigma)}{D(t+\sigma)}.$$

Alternatively,

$$\ln D(s) + \ln D(t+\sigma) \leq \ln D(s+\sigma) + \ln D(t) \tag{3.6}$$

for every $\sigma > 0$ and $s, t \in T$ with $s > t \geq 0$. Inequality (3.6) is equivalent to convexity of $\ln D(t)$.

The second part is proved analogously. Under a DU representation the second condition is equivalent to the following: for every $s, t \in T$ with $s > t \geq 0$ and every $\sigma > 0$ there exist x^* and y^* such that:

$$u(x^*)D(t+\sigma) < u(y^*)D(s+\sigma) \quad \text{but} \quad u(x^*)D(t) > u(y^*)D(s).$$

Equivalently,

$$\frac{D(s)}{D(t)}u(y^*) < u(x^*) < \frac{D(s+\sigma)}{D(t+\sigma)}u(y^*).$$

From the fact that $(y^*, s+\sigma)$ is preferred to $(x^*, t+\sigma)$ with $s > t$ we deduce that $u(y^*) > 0$. Hence,

$$\frac{D(s)}{D(t)} < \frac{D(s+\sigma)}{D(t+\sigma)}.$$

This inequality is equivalent to:

$$\ln D(s) + \ln D(t+\sigma) < \ln D(s+\sigma) + \ln D(t) \tag{3.7}$$

for every $s, t \in T$ with $s > t \geq 0$ and every $\sigma > 0$. Inequality (3.7) holds if and only if $\ln D(t)$ is *strictly* convex. □

Proposition 3.7 implies that when a DU representation exists the first condition of Definition 3.6 follows from the second one, since *strict* convexity of $\ln D(t)$ implies convexity of $\ln D(t)$. An immediate consequence is that strong present bias is equivalent to strictly DI, as stated below:

Corollary 3.8. *Suppose the preference order $\succcurlyeq$ admits a DU representation. Then it exhibits strong present bias if and only if $\succcurlyeq$ exhibits strictly DI.*

The concept of present bias requires some discussion, since there exists some inconsistency in the usage of the term in the literature. In a discrete-time setting with preferences defined over outcome streams, rather than just dated events, present bias is commonly related to quasi-hyperbolic discounting [60, 52]. Indeed, Hayashi [44] and Olea and Strzalecki [58] use a present bias axiom to provide an axiomatic foundation for quasi-hyperbolic discounting, where the present time plays a special role. Working in the same framework, Halevy [41] describes present bias as a preference reversal which may occur when an immediate outcome and a future outcome are equally delayed. However, he also calls present bias diminishing impatience.

In a continuous-time framework, Takeuchi [77] introduces a notion of present bias, defined in terms of what he calls an "equivalent delay" function, which he proves to be equivalent to DI.

Our strong present bias definition is the continuous-time analogue of Jackson and Yariv's [46, p. 4190] discrete-time present bias definition. However, there is also some inconsistency between Jackson and Yariv's present bias definition in their 2014 paper and that in their 2015 paper [47]. The continuous-time version of Jackson and Yariv's 2015 definition of present bias [47] requires the weakening of condition (ii) as follows:

(ii) for every* $s, t \in T$ *with* $s > t > 0$ *there exist* x^* *and* y^* *such that* $(x^*, t) \prec (y^*, s)$ *and* $(x^*, 0) \succ (y^*, s - t)$*.*

Condition (ii*) corresponds to Ok and Masatlioglu's definition of strong present bias [57].[3] Ok and Masatlioglu [57] introduce definitions of both present bias and strong present bias for preferences over dated outcomes in continuous time. Their definition of strong present bias is essentially a continuous-time analogue of the second condition of Jackson and Yariv's 2015 definition of present bias [47]. According to their definitions, exponential discounting exhibits present bias but not strong present bias.

It is not hard to notice that under the assumption that $\succcurlyeq$ has a DU representation, condition (ii*) corresponds to the inequality $D(s) > D(t)D(s-t)$ for any $s > t > 0$. Denoting $\sigma = s - t > 0$ we obtain $D(t+\sigma) > D(t)D(\sigma)$ for any $t, \sigma > 0$. Hence, if a DU representation is assumed, condition (ii*) is equivalent to $\ln D(t)$ being strictly superadditive for all t, i.e., $\ln D(t+\sigma) > \ln D(t) + \ln D(\sigma)$ for any $t, \sigma > 0$. This implies (by Proposition 3.7) that the continuous-time analogue of Jackson and Yariv's 2015 definition of present bias [47] requires both convexity and strict superadditivity of $\ln D(t)$, and is therefore a stronger requirement than decreasing impatience but weaker than strictly decreasing impatience. Overall the 2015 definition reflects the conventional use of the term "present bias" better as it gives a distinguished role to the present time ($t = 0$). However, since we seek an extension of Jackson and Yariv's 2014 result [46, Proposition 1] we use the definition from their 2014 paper [46], adapted to continuous time, and call it *strong* present bias to emphasise the deviation from the conventional use of the term "present bias".

3.2 Comparative DI

3.2.1 More DI and log-convexity

Assume now that there are two decision-makers and they are both decreasingly impatient. What does it mean to say that one of them is more decreasingly impatient than

[3] Ok and Masatlioglu [57] use the indifference $(x^*, t) \sim (y^*, s)$ instead of the strict preference $(x^*, t) \prec (y^*, s)$. When preferences satisfy the axioms of Fishburn and Rubinstein [32], Ok and Masatlioglu's definition of strong present bias [57] is equivalent to condition (ii*).

the other? The answer to this question is in the following definition:[4]

Definition 3.9 (cf. [63], Definition 2; [11], Definition 1)**.** *We say that* $\succcurlyeq_1$ *exhibits [strictly] more DI than* $\succcurlyeq_2$*, if for every* $\sigma > 0$*, every* ρ*, every* $s, t \in T$ *with* $0 \leq t < s$ *and every* $x, x', y, y' \in X$ *with* $y > x > 0$ *and* $y' > x' > 0$*, the conditions* $(x', t) \sim_2 (y', s)$*,* $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$ *and* $(x, t) \sim_1 (y, s)$ *imply* $(x, t + \sigma) \preccurlyeq_1 [\prec_1] (y, s + \sigma + \rho)$*.*

Not surprisingly, the (strictly) more DI relation may be expressed in terms of the comparative convexity of the logarithms of the respective discount functions, for cases in which both preference relations have DU representations.

Proposition 3.10 (cf. [63], Proposition 1)**.** *Let* $\succcurlyeq_1$ *and* $\succcurlyeq_2$ *be two preference orders with DU representations* (u_1, D_1) *and* (u_2, D_2)*, respectively. The following conditions are equivalent:*

(i) The preference order $\succcurlyeq_1$ *exhibits (strictly) more DI than* $\succcurlyeq_2$*;*

(ii) $\ln D_1(D_2^{-1}(e^z))$ *is (strictly) convex in* z *on* $(-\infty, 0]$*.*

We follow Prelec's argument for his Proposition 1 in [63]. The additional adjustment is the necessity to replace convexity of the log-transformed discount function with strict convexity for the strictly more DI case. The required adjustments are not substantial but we have included a detailed proof as it clarifies some details omitted from Prelec's original version [63]. We need to prove the following lemma first:

Lemma 3.11. *Suppose that* h_1 *and* h_2 *are strictly decreasing functions. Then* h_1 *is a (strictly) convex transformation of* h_2 *if and only if* $h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma)$ *implies that* $h_1(s) - h_1(t) \leq [<]\ h_1(s + \sigma + \rho) - h_1(t + \sigma)$ *for every* s*.* t*,* σ *and* ρ *satisfying* $0 < t < s \leq t + \sigma < s + \sigma + \rho$*.*

Proof. We prove necessity first. Suppose that h_1 is a (strictly) convex transformation of h_2; that is, there exists a (strictly) convex function f such that $h_1 = f(h_2)$. Assume also that $0 < t < s \leq t + \sigma < s + \sigma + \rho$ and

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma). \tag{3.8}$$

We need to show that

$$h_1(s) - h_1(t) \leq [<]\ h_1(s + \sigma + \rho) - h_1(t + \sigma)$$

[4] Since the sign of ρ is not restricted in Definition 3.9, it actually applies to preferences that exhibit decreasing or increasing impatience.

whenever $0 < t < s \leq t + \sigma < s + \sigma + \rho$. Since h_2 is strictly decreasing, it follows that

$$h_2(s + \sigma + \rho) < h_2(t + \sigma) \leq h_2(s) < h_2(t).$$

Recall that f is a (strictly) convex function. Therefore, as equality (3.8) holds, it implies that

$$f(h_2(t + \sigma)) - f(h_2(s + \sigma + \rho)) \leq [<]\, f(h_2(t)) - f(h_2(s)).$$

Since $h_1 = f(h_2)$, this inequality is equivalent to

$$h_1(t + \sigma) - h_1(s + \sigma + \rho) \leq [<]\, h_1(t) - h_1(s).$$

Rewriting:

$$h_1(s) - h_1(t) \leq [<]\, h_1(s + \sigma + \rho) - h_1(t + \sigma), \tag{3.9}$$

whenever $0 < t < s \leq t + \sigma < s + \sigma + \rho$.

To show the sufficiency, suppose that (3.8) implies (3.9) for every s, t, σ and ρ satisfying $0 < t < s \leq t + \sigma < s + \sigma + \rho$. Define f such that $f = h_1 \circ h_2^{-1}$. Note that we can do so because h_2^{-1} exists (since h_2 is a strictly decreasing function). Then if

$$h_2(s + \sigma + \rho) < h_2(t + \sigma) \leq h_2(s) < h_2(t)$$

and equation (3.8) holds, we have

$$f(h_2(t + \sigma)) - f(h_2(s + \sigma + \rho)) \leq [<]\, f(h_2(t)) - f(h_2(s)).$$

Therefore, f is a (strictly) convex function, which means that h_1 is a (strictly) convex transformation of h_2. □

We can now prove Proposition 3.10.

Proof. Observe that $D_i \colon [0, \infty) \to (0, 1]$ is one-to-one and onto, so $D_i^{-1} \colon (0, 1] \to [0, \infty)$.

Let us first prove that condition (i) follows from condition (ii). The proof is by contraposition. We show that not (i) implies not (ii). Assume that (i) fails; that is, there exist s and t with $0 < t < s$, $\rho > 0$, $\sigma > 0$ and $x, y, x', y' \in X$ with $0 < x < y$ and $0 < x' < y'$ such that $(x', t) \sim_2 (y', s)$, $(x', t + \sigma) \sim_2 (y', s + \sigma + \rho)$, $(x, t) \sim_1 (y, s)$ and

$$(x, t + \sigma) \succ_1 [\succcurlyeq_1]\, (y, s + \sigma + \rho).$$

Since $u_1(y) > 0$ and $u_2(y') > 0$ by assumption, this implies

$$\frac{u_2(x')}{u_2(y')} = \frac{D_2(s)}{D_2(t)} = \frac{D_2(s+\sigma+\rho)}{D_2(t+\sigma)}$$

and

$$\frac{u_1(x)}{u_1(y)} = \frac{D_1(s)}{D_1(t)} > [\geq] \frac{D_1(s+\sigma+\rho)}{D_1(t+\sigma)}.$$

Let $h_1 = \ln D_1$ and $h_2 = \ln D_2$. Note that h_1 and h_2 are both strictly decreasing functions. Observe also that $h_i \colon [0,\infty) \to (-\infty, 0]$ is one-to-one and onto. Thus $h_i^{-1} \colon (-\infty, 0] \to [0,\infty)$, where $h_i^{-1}(z) = D_i^{-1}(e^z)$. Rewriting these expressions we get $D_i(t) = e^{h_i(t)}$ for each $i \in \{1, 2\}$. Thus:

$$\frac{e^{h_2(s)}}{e^{h_2(t)}} = \frac{e^{h_2(s+\sigma+\rho)}}{e^{h_2(t+\sigma)}}$$

and

$$\frac{e^{h_1(s)}}{e^{h_1(t)}} > [\geq] \frac{e^{h_1(s+\sigma+\rho)}}{e^{h_1(t+\sigma)}}.$$

Equivalently,

$$h_2(s) - h_2(t) = h_2(s+\rho+\sigma) - h_2(t+\sigma) \tag{3.10}$$

and

$$h_1(s) - h_1(t) > [\geq]\, h_1(s+\rho+\sigma) - h_1(t+\sigma). \tag{3.11}$$

Note that $\ln D_1(D_2^{-1}(e^z))$ (strictly) convex in z on $(-\infty, 0]$ is equivalent to $h_1 \circ h_2^{-1}$ (strictly) convex in z on $(-\infty, 0]$. In other words, h_1 is a (strictly) convex transformation of h_2. By Lemma 3.11 this conclusion contradicts equation (3.10) and inequality (3.11). Therefore, not (i) implies not (ii).

Secondly, we need to demonstrate that (i) implies (ii). Using the previously introduced notation, we show that for every for every s, t, σ and ρ satisfying

$$0 < t < s \leq t + \sigma < s + \sigma + \rho$$

the equation

$$h_2(s) - h_2(t) = h_2(s+\sigma+\rho) - h_2(t+\sigma)$$

implies

$$h_1(s) - h_1(t) \leq [<]\, h_1(s+\sigma+\rho) - h_1(t+\sigma).$$

As h_1 and h_2 are decreasing functions, this proves that h_1 is a (strictly) convex transformation of h_2. Assume that $0 \leq t < s \leq t+\sigma < s+\sigma+\rho$ such that

$$h_2(s) - h_2(t) = h_2(s+\sigma+\rho) - h_2(t+\sigma).$$

By definition of $h_i = \ln D_i$ this expression is equivalent to

$$\frac{D_2(s)}{D_2(t)} = \frac{D_2(s+\sigma+\rho)}{D_2(t+\sigma)} \in (0,1).$$

As u_2 is continuous, we can choose $0 < x' < y'$ such that:

$$\frac{D_2(s)}{D_2(t)} = \frac{D_2(s+\sigma+\rho)}{D_2(t+\sigma)} = \frac{u_2(x')}{u_2(y')}.$$

Therefore, $D_2(t)u_2(x') = D_2(s)u_2(y')$ and $D_2(t+\sigma)u_2(x') = D_2(s+\sigma+\rho)u_2(y')$. This means that $(x', t) \sim_2 (y', s)$ and $(x', t+\sigma) \sim_2 (y', s+\sigma+\rho)$.

Analogously, because u_1 is continuous, we can choose x, y such that:

$$\frac{D_1(s)}{D_1(t)} = \frac{u_1(x)}{u_1(y)} \in (0,1).$$

Hence, $(x, t) \sim_1 (y, s)$.

But according to (i), if $(x', t) \sim_2 (y', s)$, $(x', t+\sigma) \sim_2 (y', s+\sigma+\rho)$ and $(x, t) \sim_1 (y, s)$ then $(x, t+\sigma) \preccurlyeq_1 [\prec_1] (y, s+\sigma+\rho)$. The latter is equivalent to:

$$\frac{D_1(s+\sigma+\rho)}{D_1(t+\sigma)} \geq [>] \frac{u_1(x)}{u_1(y)}.$$

It follows that

$$\frac{D_1(s)}{D_1(t)} \leq [<] \frac{D_1(s+\sigma+\rho)}{D_1(t+\sigma)},$$

which is equivalent to

$$\ln D_1(s) - \ln D_1(t) \leq [<] \ln D_1(s+\sigma+\rho) - \ln D_1(t+\sigma)$$

or

$$h_1(s) - h_1(t) \leq [<]\, h_1(s+\sigma+\rho) - h_1(t+\sigma).$$

Therefore,

$$h_2(s) - h_2(t) = h_2(s+\sigma+\rho) - h_2(t+\sigma)$$

implies

$$h_1(s) - h_1(t) \leq [<]\, h_1(s+\sigma+\rho) - h_1(t+\sigma)$$

whenever $0 \leq t < s \leq t+\sigma < s+\sigma+\rho$. Hence, by Lemma 3.11, h_1 is a (strictly) convex transformation of h_2. □

Note that the form of the utility functions u_1 and u_2 does not influence the comparative DI properties of preference relations.

Corollary 3.12. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ be two preference relations with DU representations (u_1, D_1) and (u_2, D_2), respectively, where $D_2(t) = \delta^t$ and $\delta \in (0, 1)$. The preference order $\succcurlyeq_1$ exhibits (strictly) DI if and only if it exhibits (strictly) more DI than $\succcurlyeq_2$.*

Proof. Prelec [63] proves that a preference relation is DI if and only if it is more DI than an exponential discount function. We prove the "strict" part of the claim.

Since $D_2(t) = \delta^t$ and $\delta \in (0, 1)$ we have

$$D_2^{-1}(e^z) \;=\; \frac{z}{\ln \delta} \;\geq\; 0.$$

By Proposition 3.10, for $\succcurlyeq_1$ to exhibit strictly more DI than $\succcurlyeq_2$ it is necessary and sufficient that $\ln D_1(D_2^{-1}(e^z))$ is strictly convex in z on $(-\infty, 0]$. However,

$$\ln D_1(D_2^{-1}(e^z)) \;=\; \ln D_1\left(\frac{z}{\ln \delta}\right) \;=\; \ln D_1(t),$$

where

$$t = \frac{z}{\ln \delta} \in [0, \infty)$$

when z takes arbitrary values in $(-\infty, 0]$. Therefore, strict convexity of $\ln D_1(D_2^{-1}(e^z))$ in z on $(-\infty, 0]$ is equivalent to strict convexity of $\ln D_1(t)$ in t on $[0, \infty)$.

By Proposition 3.5, strict convexity of $\ln D_1(t)$ in t on $[0, \infty)$ is equivalent to $\succcurlyeq_1$ exhibiting strictly DI. □

The following notations will be used below:

- If D_1 and D_2 represent equally DI preferences, we write $D_1 \sim_{DI} D_2$;
- If D_1 represents more DI preferences than D_2, we write $D_1 \succcurlyeq_{DI} D_2$;
- If D_1 represents strictly more DI preferences than D_2, we write $D_1 \succ_{DI} D_2$.

It is important to emphasize the non-standard meaning of the strict order $\succ_{DI}$ used here. As usual, if $D_1 \succ D_2$ then $D_1 \succcurlyeq D_2$ and it is not the case that $D_2 \succcurlyeq D_1$. However, the converse is false: we may have $D_1 \succcurlyeq D_2$ but neither $D_2 \succcurlyeq D_1$ nor $D_1 \succ D_2$, since there exist convex functions which are neither affine nor strictly convex.

The following corollary, due to Prelec [63], characterises the relation between any two discount functions from the same DI class.

Corollary 3.13 ([63]). *For any two discount functions D_1 and D_2, we have $D_1 \sim_{DI} D_2$ if and only if $D_1(t) = D_2(t)^c$, where $c > 0$ is a constant not depending on t.*

In fact, the "more DI" and "strictly more DI" relations are both transitive. This is established in the following proposition.

Proposition 3.14. *If $D_1 \succcurlyeq_{DI} D_2$ and $D_2 \succcurlyeq_{DI} D_3$, then $D_1 \succcurlyeq_{DI} D_3$. If at least one of the relations $D_1 \succcurlyeq_{DI} D_2$ or $D_2 \succcurlyeq_{DI} D_3$ is strict, then $D_1 \succ_{DI} D_3$.*

Proof. Suppose $D_1 \succcurlyeq_{DI} D_2$ and $D_2 \succcurlyeq_{DI} D_3$. By Proposition 3.10, we know that both $\ln D_1(D_2^{-1}(e^z))$ and $\ln D_2(D_3^{-1}(e^z))$ are convex in z on $(-\infty, 0]$. Defining $h_i = \ln D_i$ for $i \in \{1, 2, 3\}$, we can equivalently state that $h_1 \circ h_2^{-1}$ and $h_2 \circ h_3^{-1}$ are convex on $(-\infty, 0]$. To prove transitivity it is sufficient to show that $\ln D_1(D_3^{-1}(e^z))$ is convex in z on $(-\infty, 0]$, or equivalently that $h_1 \circ h_3^{-1}$ is convex on $(-\infty, 0]$.

Let $f_1 = h_1 \circ h_2^{-1}$ and $f_2 = h_2 \circ h_3^{-1}$. Then

$$\ln D_1 \left(D_3^{-1}\left(e^z\right)\right) = h_1\left(h_3^{-1}(z)\right) = h_1 h_2^{-1}\left(h_2 h_3^{-1}(z)\right) = f_1 \circ f_2(z) = f(z).$$

By the assumption, f_1 and f_2 are convex functions. Note that f_1 is increasing, as the composition of two decreasing functions h_1 and h_2^{-1}. Indeed, $h_1 = \ln D_1$ is a strictly decreasing function as D_1 is strictly decreasing, and h_2^{-1} is a decreasing function as the inverse of the decreasing function h_2. Lemma 3.1 then implies that $f(z) = f_1 \circ f_2(z) = \ln D_1\left(D_2^{-1}(e^z)\right)$ is convex, and that f is strictly convex if f_i is strictly convex for some $i \in \{1, 2\}$. □

3.2.2 Time preference rate and index of DI

In this section we assume that D is twice continuously differentiable. The *rate of time preference*, $r(t)$, is defined as follows:

$$r(t) \;=\; -\frac{D'(t)}{D(t)} = -\frac{d}{dt}\ln\left[D(t)\right].$$

The following lemma relates the DI property to the behaviour of $r(t)$.[5]

Lemma 3.15. *Let $\succcurlyeq$ be a preference relation with DU representation (u, D) in which D is twice continuously differentiable. Then the following conditions are equivalent:*

(i) The preference relation exhibits (strictly) DI;

[5] Takeuchi [77] contains a related result. His Corollary 1 says that the *hazard function* is decreasing (increasing) if and only if preferences exhibit decreasing (increasing) impatience. Takeuchi's hazard function $h(t)$ corresponds to our time preference rate $r(t)$. However, Takeuchi does not analyse the case of strictly decreasing impatience.

(ii) The time preference rate $r(t)$ is (strictly) decreasing on $[0, \infty)$.

Proof. Suppose that $r(t)$ is decreasing on $[0, \infty)$. This is equivalent to

$$r'(t) \;=\; -\frac{D''(t)D(t) - [D'(t)]^2}{D(t)^2} \;=\; \frac{[D'(t)]^2 - D''(t)D(t)}{D(t)^2} \;\leq\; 0.$$

Or, alternatively, $D''D - (D')^2 \geq 0$. This inequality is equivalent to log-convexity of D, which, by Proposition 3.5, means that the preference order exhibits DI.

To prove the equivalence of strictly DI preferences and a strictly decreasing rate of time preference, recall that a continuously differentiable function $r \colon \mathbb{R}_+ \to \mathbb{R}$ is strictly decreasing if and only if $r'(t) \leq 0$ for all t and the set $\{\, t \,|\, r'(t) = 0 \,\}$ contains no non-trivial interval [76, 69]. If a function v is differentiable on an open interval $I \subset \mathbb{R}$, then v is strictly convex on I if and only if v' is strictly increasing on I [69]. Assume that $r(t)$ is strictly decreasing on $[0, \infty)$. Let $M \subseteq \mathbb{R}_+$ be the set of t values such that $r'(t) < 0$. Then $D''(t)D(t) - [D'(t)]^2 > 0$ for all $t \in M$. Since $\mathbb{R}_+ \setminus M$ contains no non-trivial interval, $r'(t)$ being strictly decreasing is equivalent to D being strictly log-convex. □

One way to measure the level of DI for suitably differentiable discount functions was suggested by Prelec [63]. Since more DI preferences have discount functions which are more log-convex, the natural criterion would be some measure of convexity of the log of the discount function. The Arrow-Pratt coefficient, which is a measure of the concavity of a function, can be adapted to this purpose. Indeed, a non-increasing rate of time preference, $r'(t) \leq 0$, is precisely analogous to the notion of decreasing risk aversion in Pratt [61].

Recall that D is a twice continuously differentiable function. The associated *rate of impatience, $IR(D)$,* is defined as follows:

$$IR(D) \;=\; -\frac{D''}{D'}.$$

The *index of DI* of D, denoted $I_{DI}(D)$, is the difference between the rate of impatience and the rate of time preference:

$$I_{DI}(D) \;=\; IR(D) - r(D) \;=\; \left(-\frac{D''}{D'}\right) - \left(-\frac{D'}{D}\right).$$

Note that

$$I_{DI}(D)(t) \;=\; -\frac{r'(t)}{r(t)} \;=\; -\frac{d}{dt}\ln\left[r(t)\right]. \tag{3.12}$$

Prelec [63] proved that if $\succcurlyeq_1$ and $\succcurlyeq_2$ both have DU representations with twice continuously differentiable discount functions, D_1 and D_2 respectively, then $\succcurlyeq_1$ exhibits more DI than $\succcurlyeq_2$ if and only if $I_{DI}(D_1) \geq I_{DI}(D_2)$ on the interval $[0, \infty)$. The following proposition considers the "strictly more DI" case.

Proposition 3.16. *Let $\succcurlyeq_1$ and $\succcurlyeq_2$ have DU representations with discount functions D_1 and D_2, respectively, where D_1 and D_2 are twice continuously differentiable. Then the preference order $\succcurlyeq_1$ exhibits strictly more DI than $\succcurlyeq_2$ if and only if $I_{DI}(D_1) \geq I_{DI}(D_2)$ on the interval $[0, \infty)$ and $\{\, t \mid I_{DI}(D_1)(t) = I_{DI}(D_2)(t) \,\}$ contains no non-trivial interval.*

Proof. Prelec's [63, Proposition 2] proof applies the Arrow-Pratt coefficient [61], which is used to compare the concavity of functions. There is no straightforward adaptation of Prelec's argument to the case of *strict* concavity. We therefore adapt Pratt's [61] original argument directly.

Recall that D_1 is strictly more DI than D_2 if and only if $\ln(D_1)$ is strictly more convex than $\ln(D_2)$ on $[0, \infty)$. Let $h_1 = \ln(D_1)$ and $h_2 = \ln(D_2)$, so h_1 and h_2 are strictly decreasing functions. The function h_1 is strictly more convex than h_2 on $(-\infty, 0]$ if and only if there exists a strictly convex transformation f such that $h_1 = f(h_2)$, or, equivalently, $h_1\left(h_2^{-1}(z)\right)$ is strictly convex on $(-\infty, 0]$.

The first derivative of $h_1\left(h_2^{-1}(z)\right)$ is:

$$\frac{dh_1\left(h_2^{-1}(z)\right)}{dz} = \frac{h_1'\left(h_2^{-1}(z)\right)}{h_2'\left(h_2^{-1}(z)\right)}. \tag{3.13}$$

We need to show that expression (3.13) is strictly increasing. Note that $h_2^{-1}(z)$ is strictly decreasing since h_2 is strictly decreasing. Therefore, (3.13) is strictly increasing if and only if $h_1'(x)/h_2'(x)$ is strictly decreasing. The latter is satisfied if and only if

$$\log\left[\frac{h_1'(x)}{h_2'(x)}\right] \tag{3.14}$$

strictly decreases (since $\log(x)$ is strictly increasing). The first derivative of (3.14) is:

$$\frac{h_2'(x)}{h_1'(x)} \cdot \frac{h_1''(x)h_2'(x) - h_1'(x)h_2''(x)}{[h_2'(x)]^2} = \frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)}$$

Therefore (3.14) is strictly decreasing if and only if

$$\frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} \leq 0$$

and the set

$$\left\{ x \;\middle|\; \frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} = 0 \right\}$$

contains no non-trivial interval.

Note that:

$$\frac{h_i''}{h_i'} = \frac{D_i''}{D_i'} - \frac{D_i'}{D_i}.$$

Therefore,

$$\frac{h_1''(x)}{h_1'(x)} - \frac{h_2''(x)}{h_2'(x)} \leq 0$$

is equivalent to

$$-\frac{D_1''}{D_1'} - \left(-\frac{D_1'}{D_1}\right) \geq -\frac{D_2''}{D_2'} - \left(-\frac{D_2'}{D_2}\right).$$

This means that $D_1 \succ_{DI} D_2$ if and only if $I_{DI}(D_1) \geq I_{DI}(D_2)$ on $[0, \infty)$, and $\{\, t \mid I_{DI}(D_1)(t) = I_{DI}(D_2)(t) \,\}$ contains no non-trivial interval. $\square$

From Proposition 3.16, Lemma 3.15 and (3.12) it follows that $\succcurlyeq$ is DI if and only if $I_{DI}(D) \geq 0$ on $[0, \infty)$, and $\succcurlyeq$ is strictly DI if and only if $I_{DI}(D) \geq 0$ on $[0, \infty)$ and $\{\, t \mid I_{DI}(D)(t) = 0 \,\}$ contains no non-trivial interval.[6] Note that the index of DI is identically zero for an exponential discount function.

The following example illustrates the index of DI for a generalized hyperbolic discount function. We will return to this example later.

Example 3.17. *The function $D(t) = (1 + ht)^{-\alpha/h}$, with $h > 0$ and $\alpha > 0$, is called the* generalized hyperbolic *discount function. For this function we have:*

$$r(t) = \alpha(1 + ht)^{-1} \quad \textit{and} \quad IR(D)(t) = (\alpha + h)(1 + ht)^{-1}.$$

Therefore, $I_{DI}(D)(t) = h(1+ht)^{-1}$. If $D_1(t) = (1+h_1t)^{-\alpha/h_1}$ and $D_2(t) = (1+h_2t)^{-\alpha/h_2}$ are two generalized hyperbolic discount functions then $D_1 \succcurlyeq_{DI} [\succ_{DI}]\, D_2$, if and only if $h_1 \geq [>]\, h_2$.

Thus the parameter h may be used as a measure of the degree of DI of a generalized hyperbolic discount function, while the parameter α has no influence on $I_{DI}(D)$. We call parameter h the hyperbolic discount rate*. The special case of a generalized hyperbolic discount function with $\alpha = h > 0$ is called the* proportional *hyperbolic discount function.*

[6]Similarly, $\succcurlyeq$ is II if and only if $I_{DI}(D) \leq 0$ on $[0, \infty)$ and strictly II if and only if $I_{DI}(D) \leq 0$ on $[0, \infty)$ and $\{\, t \mid I_{DI}(D)(t) = 0 \,\}$ contains no non-trivial interval.

3.3 Mixtures of discount functions

As described in the introduction, there are some situations in which the necessity arises to calculate a convex combination (mixture) of discount functions.

3.3.1 Two scenarios

The first situation is a group of decision-makers with different discount functions and a social discount function needs to be constructed. One natural option is the utilitarian approach of taking a weighted average of the individual discounted utilities. This is equivalent to using the weighted average discount function when all agents have identical utility functions. This utilitarian approach has been widely used in the existing literature on time preferences ([46], [84], [82]).

A second possible scenario, discussed by Sozou [75] and Weitzman [81], arises when there is a single decision-maker with some uncertainty about her discount function. For example, there may be some possibility of not surviving to any given period, t, described by a survival function with an uncertain (constant) hazard rate [75]. Then the expected discount function of this decision-maker can be calculated as a probability-weighted average of the possible discount functions.

If the discount function D_i has probability p_i, then the expected utility of the decision-maker is

$$\hat{U}(x,t) = \sum_{i=1}^{n} p_i D_i(t) u(x) = \left(\sum_{i=1}^{n} p_i D_i(t) \right) u(x).$$

and the *certainty equivalent discount function* will be $D = \sum_{i=1}^{n} p_i D_i$.

The same question arises in both cases: Is it possible to say anything about the behaviour of the convex combination of distinct discount functions in comparison with its components, if all the component discount functions exhibit DI?

3.3.2 Mixtures of discount functions with decreasing impatience

Given a set of discount functions $\{D_1, D_2, \ldots D_n\}$, we define a *mixture* of them as

$$D = \sum_{i=1}^{n} \lambda_i D_i,$$

where $0 < \lambda_i < 1$ for all i and $\sum_{i=1}^{n} \lambda_i = 1$. Note that we define a mixture such that each D_i has a *strictly* positive weight.

We first discuss some known results on the mixture of discount functions. Two other disciplines – Risk Theory and Reliability Theory – also have relevant results that are given below. These results are rarely discussed together despite being well-known and closely related to each other.

One of the most recent results was obtained by Jackson and Yariv [46], who demonstrated that if all decision-makers in a group have exponential discount functions, but they do not all have the same discount factor, then their collective discount function must be strongly present biased (in their discrete-time setting).

It has also been noted by several authors (for example, [64] and [66]), that time preferences have strong similarities with risk preferences and that many results from risk theory are relevant in the context of intertemporal choice. The following table summarizes the similarities between risk theory and temporal discounting:

Table 3.1: Risk theory and temporal discounting

Risk theory	*Temporal discounting*
Risk aversion	Time preference rate
$RA = -\frac{u''}{u'} = -[\ln u']'$	$r = -\frac{D'}{D} = -[\ln D]'$
Decreasing risk aversion (DRA)	Decreasing time preference rate
$RA' = -[\ln u']'' \leq 0$	$r' = -[\ln D]'' \leq 0$
DRA $\Leftrightarrow$ u' is log-convex	DI $\Leftrightarrow$ D is log-convex

Pratt [61] showed that decreasing risk aversion is preserved under linear combinations. As was observed in Section 3.2.2, decreasing risk aversion is analogous to non-increasing time preference rate, or DI of the discount function.[7] Therefore, Pratt's result can be adapted to our time preference framework as follows:

Proposition 3.18 (cf. [61], Theorem 5). *Let $\succcurlyeq_1, \succcurlyeq_2, \ldots, \succcurlyeq_n$ have DU representations with twice continuously differentiable discount functions $D_1, \ldots, D_n$, respectively. Assume that $\succcurlyeq_1, \succcurlyeq_2, \ldots, \succcurlyeq_n$ all exhibit DI. Let*

$$D = \sum_{i=1}^{n} \lambda_i D_i,$$

[7] Pratt [62] remarks on the mathematical correspondence of decreasing risk aversion to decreasing time preference rate.

be a mixture of $D_1, \ldots, D_n$. Then D is DI. It is strictly DI if and only if

$$\{\, t \mid r_1(t) = r_2(t) = \ldots = r_n(t) \text{ and } r_1'(t) = r_2'(t) = \ldots = r_n'(t) = 0 \,\}$$

contains no non-trivial interval.

Proof. From the definition of time preference rate it follows that $D_i' = -r_i D_i$ for all $i = 1, \ldots, n$. The time preference rate for D is:

$$r \;=\; -\frac{D'}{D} \;=\; -\frac{\sum_{i=1}^{n} \lambda_i D_i'}{D} \;=\; \sum_{i=1}^{n} \frac{\lambda_i D_i}{D} r_i.$$

By Lemma 3.15, to prove that D exhibits DI we must show that $r'(t) \leq 0$:

$$r' \;=\; \sum_{j=1}^{n} \frac{\lambda_j D_j' \sum_{i=1}^{n} \lambda_i D_i - \lambda_j D_j \sum_{i=1}^{n} \lambda_i D_i'}{D^2} r_j + \sum_{j=1}^{n} \frac{\lambda_j D_j}{D} r_j'.$$

Rearranging and substituting $D_i' = -r_i D_i$ we obtain:

$$r' \;=\; \sum_{j=1}^{n} \frac{\lambda_j D_j}{D} r_j' + \frac{Q}{D^2},$$

where

$$Q = -\sum_{j=1}^{n} \left[\lambda_j r_j D_j \sum_{i=1}^{n} \lambda_i D_i - \lambda_j D_j \sum_{i=1}^{n} \lambda_i r_i D_i \right] r_j.$$

This is a quadratic form in $D_1, D_2, \ldots, D_n$ with the coefficient on $D_i D_j$ being

$$\lambda_i \lambda_j \left(r_i r_j - r_j^2 \right) + \lambda_i \lambda_j \left(r_i r_j - r_i^2 \right) = -\lambda_i \lambda_j (r_i - r_j)^2.$$

Hence

$$Q = -\sum_{i<j} \lambda_i \lambda_j D_i D_j (r_i - r_j)^2.$$

Since $D_i \in (0, 1]$, $\lambda_i > 0$ and $r_i' \leq 0$ for all $i = 1, \ldots n$, we have $r'(t) \leq 0$. Therefore, $\succcurlyeq$ is DI. The preference relation $\succcurlyeq$ is strictly DI if and only if $r'(t)$ is strictly decreasing. We see that $r'(t)$ is strictly decreasing iff

$$\{\, t \mid r_1(t) = r_2(t) = \ldots = r_n(t) \;\text{ and }\; r_1'(t) = r_2'(t) = \ldots = r_n'(t) = 0 \,\}$$

contains no non-trivial interval. □

Another area where results related to mixtures of discount functions can be found is Reliability Theory.[8] For example, Proschan [65] established that mixtures of distributions with decreasing failure rates always exhibit a decreasing failure rate. This result is comparable to the "non-strict" part of Proposition 3.18.

Takeuchi [77] notes that a discount function is analogous to a survival function, $S(t)$. The failure rate associated with $S(t)$ is $g(t) = -\frac{S'(t)}{S(t)}$, which behaves as a time preference rate. For twice continuously differentiable survival functions, a decreasing failure rate (DFR) corresponds to a decreasing time preference rate, and hence to DI, whereas an increasing failure rate (IFR) corresponds to II. The similarities between reliability theory and temporal discounting are given in the following table:

Table 3.2: Reliability theory and temporal discounting

Reliability theory	*Temporal discounting*
Survival function $S(t)$	Discount function $D(t)$
Failure rate (FR)	Time preference rate
$g(t) = -\frac{S'(t)}{S(t)}$	$r(t) = -\frac{D'(t)}{D(t)}$
Decreasing failure rate (DFR)	Decreasing time preference rate (DI)
$g'(t) \leq 0$	$r'(t) \leq 0$
Increasing failure rate (IFR)	Increasing time preference rate (II)
$g'(t) \geq 0$	$r'(t) \geq 0$

The following corollary describes an important special case of Proposition 3.18:

Corollary 3.19. *Mixtures of non-identical exponential discount functions are strictly DI.*

Corollary 3.19 is therefore a continuous-time version of Jackson and Yariv (2014, Proposition 1).

Prelec [63, Corollary 4] considers the mixture of two discount functions only ($n = 2$), but does not require differentiability. He proves that the mixture of two equally DI discount functions is at least as DI as each component.

Our objective is to establish a result which generalizes both Prelec [63] and Jackson and Yariv [46] (or rather, the continuous-time version in Corollary 3.19). The result we obtain is stated in the following theorem:

[8] The similarity between reliability theory and temporal discounting is discussed in [75].

Theorem 3.20. *Let $n \geq 2$ and $D_1, \dots, D_n$ be discount functions such that $D_1 \neq D_j$ for some $j > 1$ and $D_1 \sim_{DI} D_2 \sim_{DI} \dots \sim_{DI} D_n$. If D is a mixture of $D_1, \dots, D_n$, then $D \succ_{DI} D_n$.*

Proof. Since $D_1 \sim_{DI} D_2 \sim_{DI} \dots \sim_{DI} D_n$ it follows by Corollary 3.13 that $D_1(t) = D_j(t)^{c_j}$ for each $j > 1$, with $c_j > 0$ for all $j > 1$ and $c_j \neq 1$ for some $j > 1$. Therefore,

$$D = \lambda_1 D_1 + \sum_{j=2}^{n} \lambda_j D_1^{1/c_j}.$$

By Proposition 3.10 it is necessary and sufficient to demonstrate strict convexity of $f(z) = \ln D\left(D_1^{-1}(e^z)\right)$ for $z \leq 0$. We note that:

$$f(z) = \ln D\left(D_1^{-1}(e^z)\right) = \ln\left(\lambda_1 e^z + \sum_{j=2}^{n} \lambda_j e^{z/c_j}\right) = \ln\left(\sum_{i=1}^{n} \lambda_i e^{z/c_i}\right),$$

where $c_1 = 1$. The first-order derivative of $f(z)$ is:

$$f'(z) = \frac{\sum_{i=1}^{n} \lambda_i \frac{1}{c_i} e^{z/c_i}}{\sum_{i=1}^{n} \lambda_i e^{z/c_i}}.$$

The second-order derivative is:

$$f''(z) = \frac{\left(\sum_{i=1}^{n} \lambda_i \left(\frac{1}{c_i}\right)^2 e^{z/c_i}\right)\left(\sum_{i=1}^{n} \lambda_i e^{z/c_i}\right) - \left(\sum_{i=1}^{n} \lambda_i \frac{1}{c_i} e^{z/c_i}\right)^2}{\left[\sum_{i=1}^{n} \lambda_i e^{z/c_i}\right]^2} = \frac{p(z)}{q(z)}.$$

The denominator $q(z)$ of this fraction is strictly positive. We can simplify the numerator $p(z)$ as follows:

$$p(z) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j e^{z/c_i} e^{z/c_j} \left(\frac{1}{c_i}\right)^2 - \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j e^{z/c_i} e^{z/c_j} \frac{1}{c_i}\frac{1}{c_j}.$$

Therefore, we have:

$$p(z) = \sum_{i=1}^{n}\sum_{j=1}^{n} \theta_{ij} \frac{1}{c_i}\left(\frac{1}{c_i} - \frac{1}{c_j}\right)$$

where $\theta_{ij} = \lambda_i \lambda_j e^{z/c_i} e^{z/c_j}$. Since $\theta_{ij} = \theta_{ji} > 0$ for all i and j we see that

$$p(z) = \sum_{i<j} \theta_{ij}\left[\frac{1}{c_i}\left(\frac{1}{c_i} - \frac{1}{c_j}\right) + \frac{1}{c_j}\left(\frac{1}{c_j} - \frac{1}{c_i}\right)\right] = \sum_{i<j} \theta_{ij}\left(\frac{1}{c_i} - \frac{1}{c_j}\right)^2,$$

where $c_1 = 1$. Hence, $p(z) > 0$ since $c_j \neq c_1$ for some $j > 1$.

Therefore, $f(z)$ is a strictly convex function. $\square$

Given Theorem 3.20, it is natural to suppose that if $D_1 \succcurlyeq_{DI} D_2 \succcurlyeq_{DI} \cdots \succcurlyeq_{DI} D_n$ then $D \succ_{DI} D_n$ for any mixture of $D_1, D_2, \ldots, D_n$. Surprisingly, however, this need not be the case. The following example illustrates.

Example 3.21. *Let $\rho_1 > \rho_2 > 0$ and let $\bar{t} > 0$. Define $D_2(t) = \exp(-\rho_2 t)$ and*

$$D_1(t) = \begin{cases} \exp(-\rho_1 t) & \text{if } t < \bar{t} \\ \exp\left(-\rho_2 t - \bar{d}\right) & \text{if } t \geq \bar{t} \end{cases}$$

where $\bar{d} = (\rho_1 - \rho_2)\bar{t} > 0$. It is a continuous-time analogue of a quasi-hyperbolic discount function. Observe that

$$D_2^{-1}(e^z) = -\frac{z}{\rho_2}$$

and therefore

$$\ln D_1\left(D_2^{-1}(e^z)\right) = \begin{cases} (\rho_1/\rho_2)\, z & \text{if } z > -\rho_2\bar{t} \\ z - \bar{d} & \text{if } z \leq -\rho_2\bar{t} \end{cases} \tag{3.15}$$

Since

$$\frac{\rho_1}{\rho_2}\left(-\rho_2\bar{t}\right) = -\rho_1\bar{t} = \left(-\rho_2\bar{t}\right) - \bar{d},$$

(3.15) is continuous. From this fact and $(\rho_1/\rho_2) > 1$ we deduce that (3.15) is convex, but neither affine nor strictly convex. It follows that $D_1 \succcurlyeq_{DI} D_2$.

Let $D = \lambda D_1 + (1-\lambda) D_2$ with $0 < \lambda < 1$. Then

$$D(t) = \begin{cases} \lambda \exp(-\rho_1 t) + (1-\lambda)\exp(-\rho_2 t) & \text{if } t < \bar{t} \\ \left[\lambda \exp(-\bar{d}) + (1-\lambda)\right] \exp(-\rho_2 t) & \text{if } t \geq \bar{t} \end{cases}$$

and therefore, for any $t \geq \bar{t}$,

$$\begin{aligned} \ln(D(t)) &= \ln\left[\lambda \exp(-\bar{d}) + (1-\lambda)\right] - \rho_2 t \\ &= \ln\left[\lambda \exp(-\bar{d}) + (1-\lambda)\right] + \ln(D_2(t)). \end{aligned}$$

*In other words, $\ln D$ is an affine transformation of $\ln D_2$ on $\left[\bar{t}, \infty\right)$. In particular, it is **not** the case that $D \succ_{DI} D_2$.*

Example 3.21 shows that mixing DI-ordered discount functions need not produce a mixture that is strictly more DI than the "least DI" of its components. However, the following weaker result can be established.

Theorem 3.22. *Let $n \geq 2$ and let $D_1, D_2, \ldots, D_n$ be discount functions such that $D_1 \succcurlyeq_{DI} D_2 \succcurlyeq_{DI} \ldots \succcurlyeq_{DI} D_n$. If D is a mixture of $D_1, D_2, \ldots, D_n$, then $D \succcurlyeq_{DI} D_n$. Moreover, if $D_i \succ_{DI} D_{i+1}$ for some $i \in \{1, 2, ..., n-1\}$ then $D \succ_{DI} D_n$.*

Proof. We argue by induction. Let $n = 2$. Then

$$D\left(D_2^{-1}(e^z)\right) \;=\; \lambda_1 D_1\left(D_2^{-1}(e^z)\right) + \lambda_2 e^z. \tag{3.16}$$

Since $D_1 \succcurlyeq_{DI} D_2$ we know that $D_1\left(D_2^{-1}(e^z)\right)$ is log-convex, and hence so is $\lambda_1 D_1\left(D_2^{-1}(e^z)\right)$. Since $\lambda_2 e^z$ is obviously log-convex, it follows by a result of Artin ([10]) that the sum (3.16) is log-convex. That is, $D \succcurlyeq_{DI} D_2$. If $D_1 \succ_{DI} D_2$ then (3.16) is strictly log-convex by Lemma 3.2. Hence, $D \succ_{DI} D_2$.

Suppose $k \geq 2$ and the result is true for all $n \leq k$. Let $n = k+1$. We have:

$$D \;=\; \sum_{i=1}^{k+1} \lambda_i D_i \;=\; (1-\lambda_{k+1})\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] + \lambda_{k+1} D_{k+1}$$

The inductive hypothesis implies that

$$\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] \succcurlyeq_{DI} D_k \tag{3.17}$$

and hence

$$\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] \succcurlyeq_{DI} D_{k+1}$$

by Proposition 3.14. Therefore, applying the inductive hypothesis once more, we deduce that $D \succcurlyeq_{DI} D_{k+1}$. If $D_k \succ_{DI} D_{k+1}$, then (3.17) and Proposition 3.14 imply $D \succ_{DI} D_{k+1}$. If $D_i \succ_{DI} D_{i+1}$ for some $i \in \{1, 2, ..., k-1\}$ then

$$\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] \succ_{DI} D_k$$

by the inductive hypothesis, and hence

$$\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] \succ_{DI} D_{k+1}$$

by Proposition 3.14. Applying the inductive hypothesis once more, we have $D \succ_{DI} D_{k+1}$. □

3.3.3 Mixtures of twice continuously differentiable discount functions

Note that when discount functions in Theorem 3.22 are twice continuously differentiable, Theorem 3.22 and Proposition 3.16 imply that

$$I_{DI}(D) \geq \min_i \{I_{DI}(D_i)\} \quad \text{on} \quad [0, \infty) \tag{3.18}$$

and the set of t values at which equality holds does not include any non-trivial interval.

The next two examples illustrate this fact.

Example 3.23. *Let $D_i(t) = \exp(-r_i t)$, with $r_1 = 0.01, r_2 = 0.02, r_3 = 0.03$. Consider their mixture $D = 1/3(D_1 + D_2 + D_3)$. The Index of DI for each discount function is $I_{DI}(D_i(t)) = 0$ for all t, whereas the Index of DI for the mixture is:*

$$I_{DI}(D) = \frac{e^{-(r_1+r_2)t}(r_1 - r_2)^2 + e^{-(r_2+r_3)t}(r_2 - r_3)^2 + e^{-(r_1+r_3)t}(r_1 - r_3)^2}{(r_1 e^{-r_1 t} + r_2 e^{-r_2 t} + r_3 e^{-r_3 t})(e^{-r_1 t} + e^{-r_2 t} + e^{-r_3 t})}.$$

Clearly, this is strictly greater than zero everywhere, as also shown on Figure 3.1.

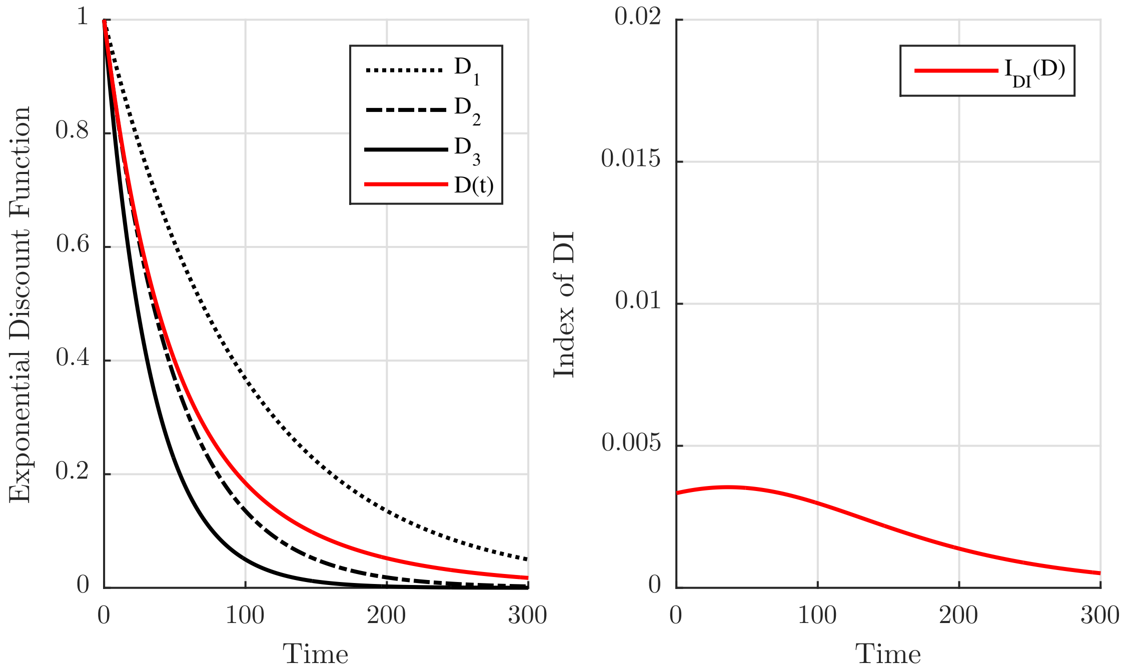

Figure 3.1: Index of DI for the mixture of exponential discount functions

Example 3.24. *Let $D_i(t) = (1 + ht)^{-a_i/h}$, with $a_1 = 1, a_2 = 2, a_3 = 3, h = 5$. Consider their mixture $D = 1/3(D_1 + D_2 + D_3)$. Recall from Example 3.17 that $I_{DI}(D_i)(t) = h_i(1 + h_i t)^{-1} = 5(1 + 5t)^{-1}$. From Figure 3.2 it can be seen that $I_{DI}(D)$ lies strictly above $I_{DI}(D_i)$.*

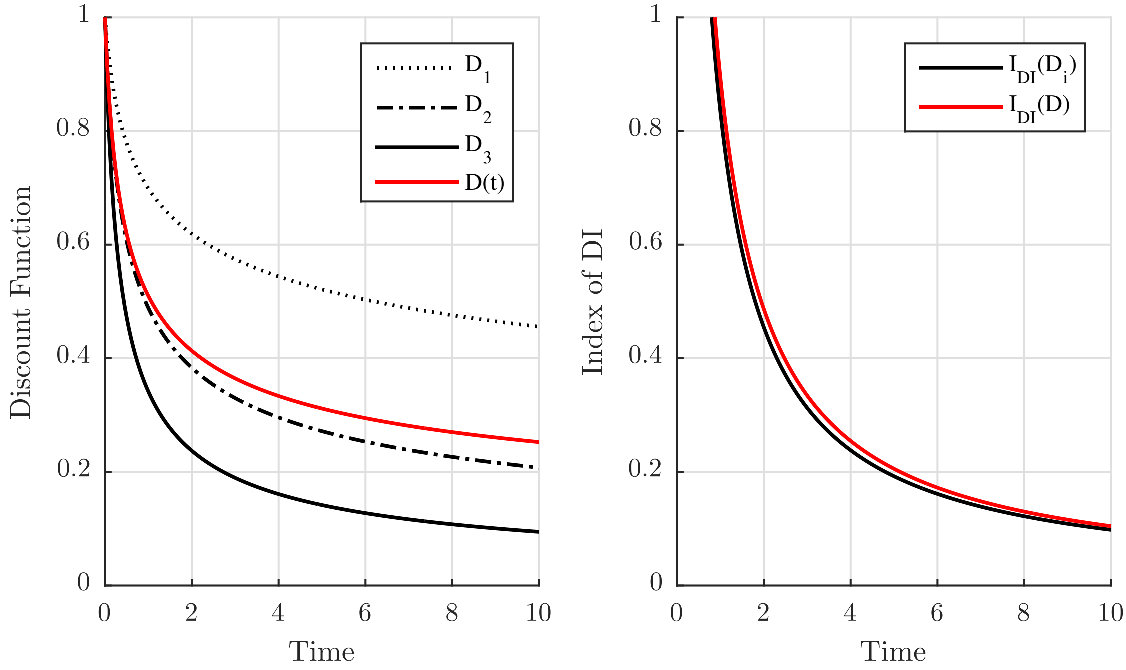

Figure 3.2: Index of DI for the mixture of generalized hyperbolic discount functions

While Theorem 3.20 only guarantees (3.18) when the D_i functions are DI-comparable, one can show that (3.18) holds more generally.

Theorem 3.25. *Let $\succcurlyeq_1, \succcurlyeq_2, \ldots, \succcurlyeq_n$ have DU representations with twice continuously differentiable discount functions $D_1, D_2, \ldots, D_n$, respectively. Let $D = \sum_{i=1}^n \lambda_i D_i$ be a mixture of $D_1, D_2, \ldots, D_n$. Then $I_{DI}(D) \geq \min_i\{I_{DI}(D_i)\}$ on $[0, \infty)$, and*

$$I_{DI}(D)(\hat{t}) > \min_i\{I_{DI}(D_i)(\hat{t})\}$$

if $r_j(\hat{t}) \neq r_k(\hat{t})$ for some $j \neq k$.

Proof. Let $I_i = I_{DI}(D_i)$ for all $i \in \{1, \ldots, n\}$ and let $I = I_{DI}(D)$. Recall that $D' = -rD$ and hence $D'' = Dr^2 - Dr' = Dr(r + I)$. Recall also that

$$I \;=\; -\frac{D''}{D'} + \frac{D'}{D}.$$

Therefore,

$$I \;=\; \frac{-\sum_{i=1}^n \lambda_i D_i''}{\sum_{i=1}^n \lambda_i D_i'} + \frac{\sum_{i=1}^n \lambda_i D_i'}{\sum_{i=1}^n \lambda_i D_i} \;=\; \frac{\sum_{i=1}^n \lambda_i D_i r_i (r_i + I_i)}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i}.$$

This expression can be rearranged as follows:

$$I = \frac{\sum_{i=1}^n \lambda_i D_i r_i I_i}{\sum_{i=1}^n \lambda_i D_i r_i} + \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i} = \sum_{i=1}^n \alpha_i(t) I_i + Q,$$

where

$$Q = \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i} \quad \text{and} \quad \alpha_i = \frac{\lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i r_i}$$

with $\sum_{i=1}^n \alpha_i = 1$ and $\alpha_i \geq 0$. Note that

$$\min_i\{I_i\} \leq \sum_{i=1}^n \alpha_i I_i \leq \max_i\{I_i\} \text{ for all } t \in [0, \infty).$$

The expression Q can be rewritten as:

$$Q = \frac{\left[\sum_{i=1}^n \lambda_i D_i r_i^2\right] \cdot \left[\sum_{i=1}^n \lambda_i D_i\right] - \left[\sum_{i=1}^n \lambda_i D_i r_i\right]^2}{\left[\sum_{i=1}^n \lambda_i D_i r_i\right] \cdot \left[\sum_{i=1}^n \lambda_i D_i\right]}.$$

The denominator of Q is strictly positive, so the sign of Q depends on the sign of the numerator. Let N be the numerator of Q:

$$N = \left[\sum_{i=1}^n \lambda_i D_i r_i^2\right] \cdot \left[\sum_{i=1}^n \lambda_i D_i\right] - \left[\sum_{i=1}^n \lambda_i D_i r_i\right]^2.$$

We can simplify N as follows:

$$N = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j D_i D_j r_i^2 - \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j D_i D_j r_i r_j.$$

Therefore, we have:

$$N = \sum_{i=1}^n \sum_{j=1}^n \theta_{ij} r_i (r_i - r_j)$$

where $\theta_{ij} = \lambda_i \lambda_j D_i D_j$. Since $\theta_{ij} = \theta_{ji} > 0$ for all i and j we see that

$$N = \sum_{i<j} \theta_{ij} \left[r_i (r_i - r_j) + r_j (r_j - r_i)\right] = \sum_{i<j} \theta_{ij} (r_i - r_j)^2.$$

Hence, $N \geq 0$ and $N > 0$ if $r_j \neq r_k$ for some $j \neq k$. It follows that $Q \geq 0$ and $Q > 0$ if $r_j \neq r_k$ for some $j \neq k$. Therefore, since $I = \sum_{i=1}^n \alpha_i I_i + Q$ and

$$\min_i\{I_i\} \leq \sum_{i=1}^n \alpha_i I_i \leq \max_i\{I_i\} \text{ for all } t \in [0, \infty),$$

we have:

$$\min_i\{I_i\} \;\le\; \min_i\{I_i\} + Q \;\le\; \sum_{i=1}^{n} \alpha_i I_i + Q = I.$$

In other words, $I \ge \min_i\{I_i\}$ on $[0,\infty)$, and $I(\hat{t}) > \min_i\{I_i(\hat{t})\}$ if $r_j(\hat{t}) \ne r_k(\hat{t})$ for some $j \ne k$. □

Observe that this result does not require the discount functions to exhibit decreasing impatience. The following example provides an illustration of Theorem 3.25.

Example 3.26. *Let $D_1(t) = (1+ht)^{-2}$ be a zero-speed hyperbolic discount function [48] and $D_2(t) = \exp\left(-\alpha t^{1/2}\right)$ be a slow Weibull discount function [48]. As shown in Example 3.17, $I_{DI}(D_1)(t) = h(1+ht)^{-1} > 0$ for all t. We also have $I_{DI}(D_2)(t) = (2t)^{-1} > 0$ for all t since*

$$r_2(t) = \frac{\alpha}{2}t^{-1/2} \quad \textit{and} \quad r_2'(t) = -\frac{\alpha}{4}t^{-3/2}.$$

Therefore, both D_1 and D_2 exhibit strict DI. Assume that $h = 0.1$. Then

$$I_{DI}(D_1)(t) - I_{DI}(D_2)(t) \;=\; \frac{0.1}{1+0.1t} - 0.5\frac{1}{t} \;=\; 0.05\frac{t-10}{t(1+0.1t)}.$$

Obviously, $I_{DI}(D_1)(t) \le I_{DI}(D_2)(t)$ if and only if $t \le 10$ and $I_{DI}(D_1)(t) > I_{DI}(D_2)(t)$ if and only if $t > 10$. It follows that D_1 and D_2 both are from incomparable DI classes. Since D_1 and D_2 both exhibit strictly DI, Proposition 3.18 implies that their mixture D also exhibits strictly DI.

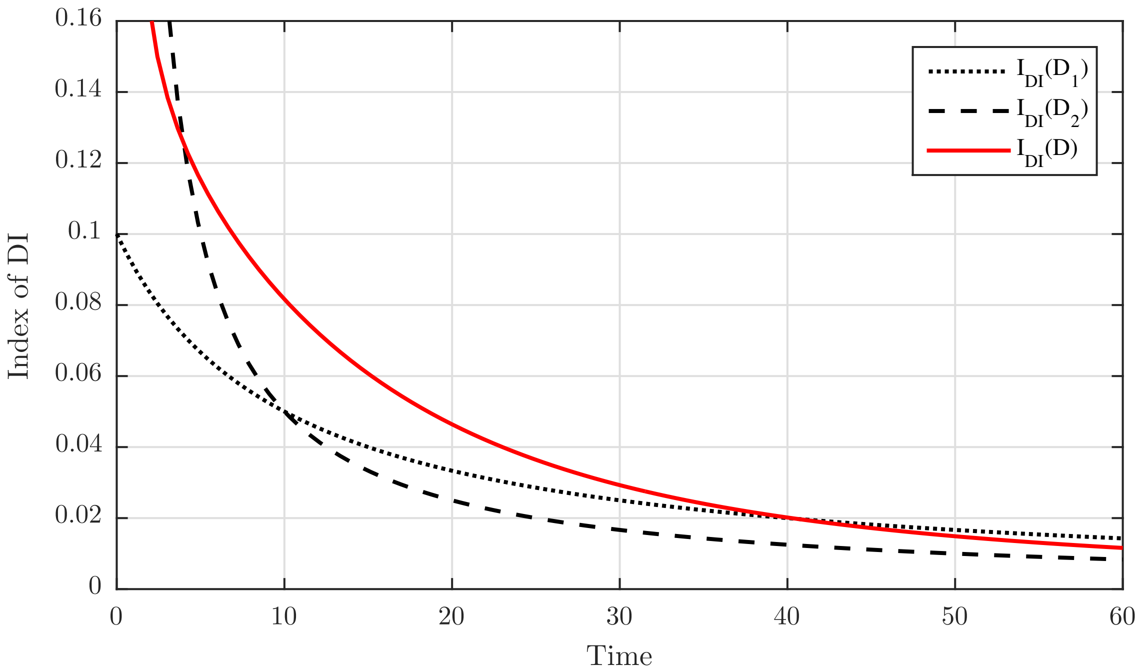

Figure 3.3: Index of DI for the mixture of D_1 and D_2 from incomparable DI classes

By direct calculation we obtain the following expression:

$$I_{DI}(D)(t) = \frac{6\lambda_1 h^2(1+ht)^{-4} + (1/4)\lambda_2\alpha \exp -\alpha t^{0.5} t^{-1}(\alpha + t^{-0.5})}{2\lambda_1 h(1+ht)^{-3} + (1/2)\lambda_2 \alpha \exp -\alpha t^{0.5} t^{-0.5}}$$

$$- \frac{2\lambda_1 h(1+ht)^{-3} + (1/2)\lambda_2\alpha \exp -\alpha t^{0.5} t^{-0.5}}{\lambda_1(1+ht)^{-2} + \lambda_2 \exp -\alpha t^{0.5}}.$$

The behaviour of $I_{DI}(D)$ with parameters $\lambda_1 = \lambda_2 = 0.5$, $h = 0.1$ and $\alpha = 0.12$ is illustrated in Figure 3.3. It can be clearly seen from Figure 3.3 that neither $D \succ_{DI} D_1$ nor $D \succ_{DI} D_2$. However, $I_{DI}(D) \geq \min\{I_{DI}(D_1), I_{DI}(D_2)\}$ on $[0, \infty)$.

3.3.4 Mixtures of proportional hyperbolic discount functions

Weitzman [81] shows that if the appropriate discount function is uncertain, then future costs and benefits must eventually be discounted at the lowest possible limiting time preference rate. This result is particularly salient when the possible discount functions are all exponential, with constant time preference rates. The purpose of this section is to give an analogous result for proportional hyperbolic discount functions, with constant hyperbolic discount rates (Example 3.17). The result in this case is very different to Weitzman's. Long-term future benefits and costs are discounted, not at the lowest hyperbolic discount rate, but at the probability-weighted harmonic mean of the individual hyperbolic discount rates.

Suppose that there is some uncertainty about the rate of time preference, and we have a set of possible scenarios $N = \{1, \ldots, n\}$ where time preference rate $r_i(t)$ has probability $p_i \geq 0$, such that $\sum_{t=1}^{n} p_i = 1$. Since

$$r_i(t) = -\frac{D_i'(t)}{D_i(t)},$$

the corresponding discount function can be expressed in terms of the rate of time preference as follows

$$D_i(t) = \exp\left(-\int_0^t r_i(\tau)d\tau\right) \text{ for each } i \in N. \tag{3.19}$$

The certainty equivalent discount function will be:

$$D = \sum_{i=1}^{n} p_i D_i, \text{ where } p_i \geq 0 \text{ and } \sum_{t=1}^{n} p_i = 1.$$

Then the certainty equivalent time preference rate is $r = -\frac{D'}{D}$. Weitzman [81] proved that if each $r_i(t)$ converges to a non-negative value as time goes to infinity, then the certainty equivalent rate of time preference converges to the lowest of these limit values. In other words, if $\lim_{t\to\infty} r_i(t) = r_i^*$, then $\lim_{t\to\infty} r(t) = \min\{r_1^*, \ldots, r_n^*\}$.

Example 3.27. *Note that $r_i(t)$ in (3.19) is constant if and only if D_i is exponential. In this case we have:*

$$D_i(t) = \exp(-r_i t) \text{ for each } i \in N,$$

where $r_i = const$. Therefore, Weitzman's result implies that $\lim_{t\to\infty} r(t) = \min_i r_i$. Figure 3.4 illustrates for the case $n = 3$, $r_1 = 0.01$, $r_2 = 0.02$, $r_3 = 0.03$ and $p_1 = p_2 = p_3 = 1/3$. We also observe that the certainty equivalent rate of time preference $r(t)$ decreases monotonically towards r_1. This is a consequence of Corollaries 3.12 and 3.19 and the fact that $I_{DI}(D) = -r'/r$.

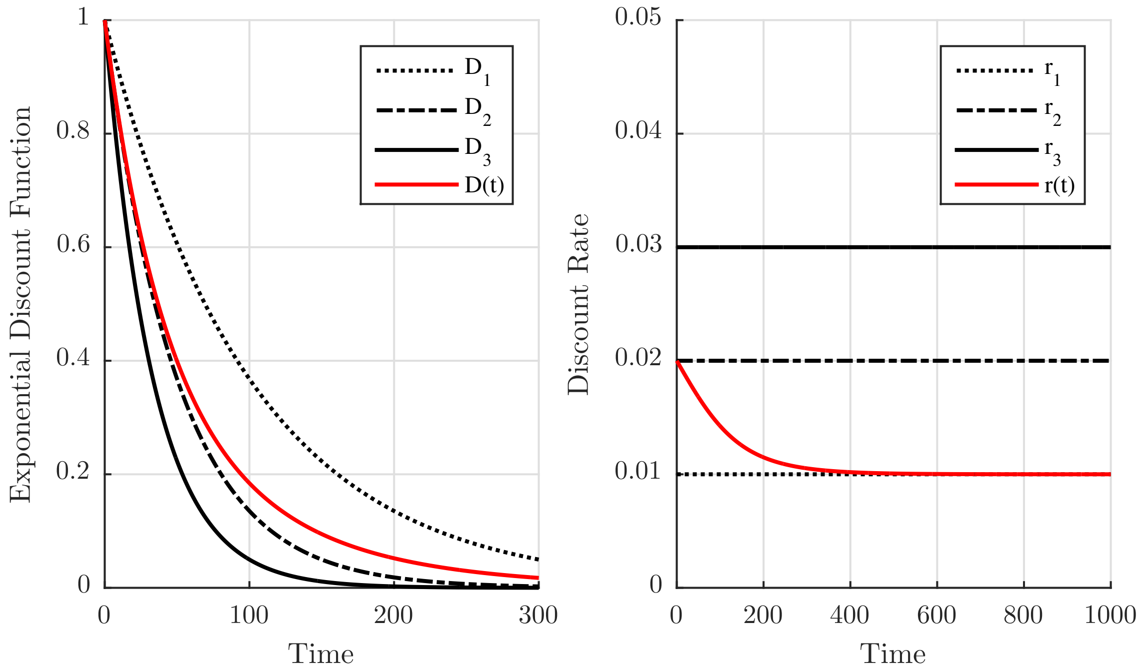

Figure 3.4: Mixture of exponential discount functions and its associated discount rate

However, Weitzman's result [81] does not provide much insight in the special case when each possible time preference has a DU representation with a proportional hyperbolic discount function. Suppose

$$D_i(t) = \frac{1}{1 + h_i t}$$

for each $i \in N$, where $h_i > 0$ is the hyperbolic discount rate. Without loss of generality we assume that $h_1 > h_2 > \ldots > h_n$. Suppose that D_i eventuates with probability p_i where $p_i \geq 0$ and $\sum_{i=1}^{n} p_i = 1$. Then the certainty equivalent discount function would

be
$$D(t) \;=\; \frac{p_1}{1+h_1 t} + \ldots + \frac{p_n}{1+h_n t}.$$
The rate of time preference is
$$r_i(t) \;=\; \frac{h_i}{1+h_i t}$$
for all i. It is obvious that $r_i^* = r_j^* = 0$ for all $i \neq j$ and $\lim_{t\to\infty} r(t) = 0$, which, indeed, corresponds to Weitzman's result. However, this conclusion does not give much information about the asymptotic behaviour of the certainty equivalent discount function. Given that each possible discount function comes from a different DI class (unlike in the case of heterogeneous exponential discount functions) we would like to know which (if any) most closely characterises the asymptotic behaviour of the certain equivalent function.

To answer this question we need to modify the analysis of Weitzman. Note that the certainty equivalent discount function can be written as
$$D(t) \;=\; \frac{1}{1+h(t)t},$$
where $h(t)$ is the *certainty equivalent hyperbolic discount rate*. In particular,
$$h(t) = \left(\frac{1}{D(t)} - 1\right)\frac{1}{t},$$
so $h(t)$ is well-defined for $t \in (0,\infty)$. We ask: *How does $h(t)$ behave as $t \to \infty$?*

We remind the reader that the *weighted harmonic mean* of non-negative values $x_1, x_2, \ldots, x_n$ with non-negative weights $a_1, a_2, \ldots, a_n$ satisfying $a_1 + \ldots + a_n = 1$ is
$$H(x_1, a_1; \ldots; x_n, a_n) = \left(\sum_{i=1}^{n} \frac{a_i}{x_i}\right)^{-1}.$$
It is well-known that the weighted harmonic mean is smaller than the corresponding expected value (weighted arithmetic mean).

Theorem 3.28. *Suppose that each D_i ($i \in N$) is a proportional hyperbolic discount function, with associated hyperbolic discount rate h_i. Discount function D_i has probability p_i. Then the long-term certainty equivalent hyperbolic discount rate is the probability-weighted harmonic mean of the individual hyperbolic discount rates, $H(h_1, p_1; \ldots; h_n, p_n)$.*

Proof. We note that
$$\frac{p_i}{1+h_i t} \;=\; \frac{p_i}{h_i t} + \epsilon_i(t),$$

where $\epsilon_i(t)/t^2 \to 0$ when $t \to \infty$. Let $\epsilon(t) = \epsilon_1(t) + \ldots + \epsilon_n(t)$. Hence it follows that:

$$\begin{aligned}
\frac{1}{1+h(t)t} = \sum_{i=1}^{n} p_i D_i(t) &= \frac{p_1}{1+h_1 t} + \ldots + \frac{p_n}{1+h_n t} \\
&= \frac{p_1}{h_1 t} + \ldots + \frac{p_n}{h_n t} + \epsilon(t) \\
&= \left(\frac{p_1}{h_1} + \ldots + \frac{p_n}{h_n}\right)\frac{1}{t} + \epsilon(t) \\
&= \frac{1}{H(h_1, p_1; \ldots; h_n, p_n)t} + \epsilon(t) \\
&= \frac{1}{1 + H(h_1, p_1; \ldots; h_n, p_n)t} + \hat{\epsilon}(t),
\end{aligned}$$

where $\hat{\epsilon}(t)/t^2 \to 0$ as $t \to \infty$. This implies that $h(t) \to H(h_1, p_1; \ldots; h_n, p_n)$ as $t \to \infty$. $\square$

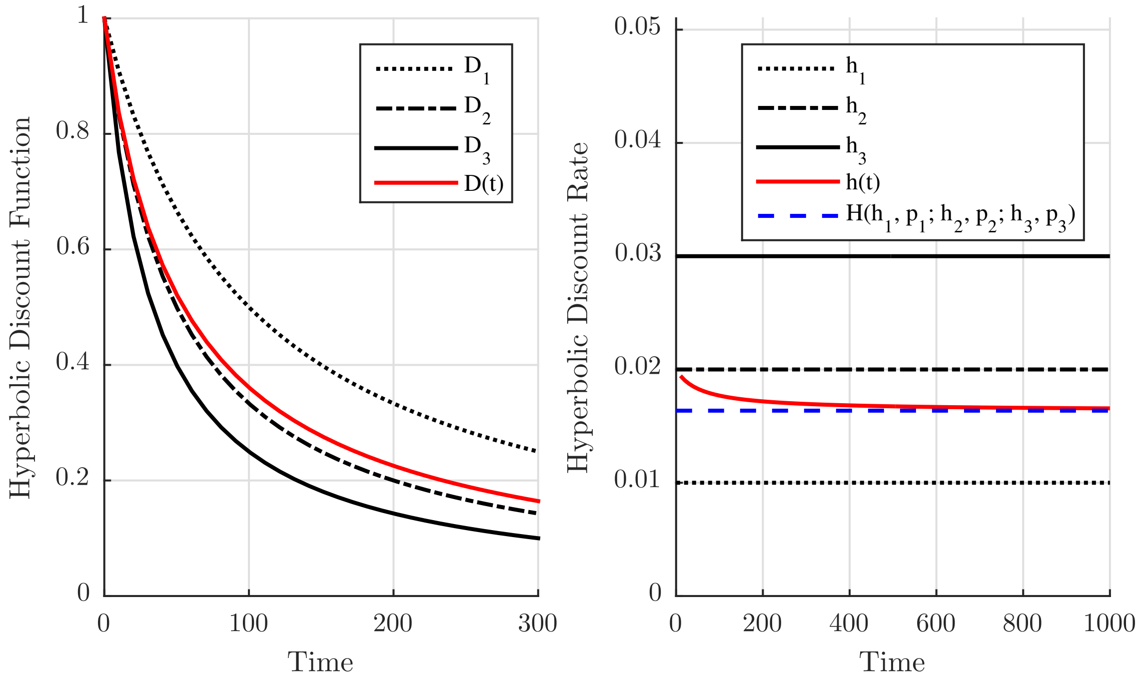

Figure 3.5: Mixture of hyperbolic discount functions and its associated hyperbolic discount rate

Figure 3.5 illustrates Theorem 3.28 for the case $n = 3$, when hyperbolic rates $h_1 = 0.01$, $h_2 = 0.02$ and $h_3 = 0.03$ occur with equal probabilities. Note that $h_2 = 0.02$ corresponds to the arithmetic mean of h_1, h_2 and h_3. Figure 3.5 displays the convergence of the certainty equivalent hyperbolic discount rate to the weighted harmonic mean $H(h_1, p_1; h_2, p_;h_3, p_3)$. It also shows the certainty equivalent hyperbolic discount rate decreasing monotonically. The following proposition proves that this is always the case.

Proposition 3.29. *Suppose that each D_i ($i \in N$) is a proportional hyperbolic discount function, with associated hyperbolic discount rate h_i. Discount function D_i will eventuate with probability p_i. Then the certainty equivalent hyperbolic discount rate is strictly decreasing on* $(0, \infty)$.

Proof. We prove this statement by induction on n. First we need to prove that the statement holds for $n = 2$. The respective certainty equivalent hyperbolic discount rate is:

$$h(t) = \left[\frac{1}{p_1(1+h_1t)^{-1} + p_2(1+h_2t)^{-1}} - 1\right]\frac{1}{t}$$

for each $t > 0$. Rearranging:

$$h(t) = \left[\frac{(1+h_1t)(1+h_2t)}{p_1(1+h_2t) + p_2(1+h_1t)} - 1\right]\frac{1}{t} = \left[\frac{1+(h_1+h_2)t + h_1h_2t^2}{p_1+p_2+(p_1h_2+p_2h_1)t} - 1\right]\frac{1}{t}.$$

Since $p_1 + p_2 = 1$ we obtain:

$$\begin{aligned} h(t) = \left[\frac{1+(h_1+h_2)\,t + h_1h_2t^2}{1+(p_1h_2+p_2h_1)\,t} - 1\right]\frac{1}{t} &= \frac{(h_1+h_2-p_1h_2-p_2h_1)\,t + h_1h_2t^2}{1+(p_1h_2+p_2h_1)\,t}\cdot\frac{1}{t} \\ = \frac{h_1+h_2-p_1h_2-p_2h_1+h_1h_2t}{1+(p_1h_2+p_2h_1)\,t} &= \frac{p_1h_1+p_2h_2+h_1h_2t}{1+(p_1h_2+p_2h_1)\,t}. \end{aligned}$$

By differentiating $h(t)$:

$$h'(t) = \frac{h_1h_2\left(1+(p_1h_2+p_2h_1)\,t\right) - (p_1h_1+p_2h_2+h_1h_2t)\,(p_1h_2+p_2h_1)}{\left[1+(p_1h_2+p_2h_1)t\right]^2} \tag{3.20}$$

We need to show that $h'(t) < 0$. Since the denominator of (3.20) is positive, the sign of $h'(t)$ depends on the sign of the numerator. Therefore, we denote the numerator of (3.20) by Q and analyse it separately:

$$\begin{aligned} Q(t) &= h_1h_2\left[1+(p_1h_2+p_2h_1)\,t\right] - (p_1h_1+p_2h_2+h_1h_2t)\,(p_1h_2+p_2h_1) \\ &= h_1h_2 + h_1h_2(p_1h_2+p_2h_1)t - (p_1h_1+p_2h_2)(p_1h_2+p_2h_1) - h_1h_2(p_1h_2+p_2h_1)t \\ &= h_1h_2 - (p_1h_1+p_2h_2)\,(p_1h_2+p_2h_1)\,. \end{aligned}$$

By expanding the brackets and using the fact that $p_1+p_2 = 1$ implies $1-p_1^2-p_2^2 = 2p_1p_2$ expression Q can be simplified further:

$$\begin{aligned} Q(t) &= h_1h_2 - p_1^2h_1h_2 - p_1p_2h_1^2 - p_1p_2h_2^2 - p_2^2h_1h_2 \\ &= h_1h_2\left(1-p_1^2-p_2^2\right) - p_1p_2\left(h_1^2+h_2^2\right) \\ &= 2p_1p_2h_1h_2 - p_1p_2\left(h_1^2+h_2^2\right) \end{aligned}$$

$$= -p_1 p_2 (h_1 - h_2)^2 .$$

Therefore, since $h_1 \neq h_2$ we have $Q < 0$. Hence it follows that $h'(t) < 0$ and $h(t)$ is strictly decreasing.

Suppose that the proposition holds for $n = k$. We need to show that it also holds for $n = k + 1$. When $n = k + 1$ the certainty equivalent hyperbolic discount rate is:

$$h_{k+1}(t) = \left[\frac{1}{D^{(k+1)}} - 1 \right] \frac{1}{t},$$

where

$$D^{(k+1)} = \sum_{i=1}^{k+1} p_i D_i = (1 - p_{k+1}) \left(\sum_{i=1}^{k} \frac{p_i}{1 - p_{k+1}} D_i \right) + p_{k+1} D_{k+1}.$$

Since

$$\sum_{i=1}^{k} \frac{p_i}{1 - p_{k+1}} = 1,$$

we have

$$D^{(k+1)} = (1 - p_{k+1}) D^{(k)} + p_{k+1} D_{k+1}.$$

where

$$D^{(k)} = \sum_{i=1}^{k} \frac{p_i}{1 - p_{k+1}} D_i.$$

By the induction hypothesis it follows that

$$D^{(k)} = \frac{1}{1 + h_k(t)t},$$

where h_k is strictly decreasing. Therefore,

$$\begin{aligned} h^{(k+1)}(t) &= \left[\frac{1}{(1 - p_{k+1}) D^{(k)} + p_{k+1} D_{k+1}} - 1 \right] \frac{1}{t} \\ &= \left[\frac{1}{(1 - p_{k+1}) (1 + h_k(t)t)^{-1} + p_{k+1} (1 + h_{k+1} t)^{-1}} - 1 \right] \frac{1}{t}. \end{aligned}$$

Let $\hat{p_1} = 1 - p_{k+1}$, $\hat{p_2} = p_{k+1}$, $\hat{h_1}(t) = h_k(t)$ and $\hat{h_2} = h_{k+1} = const$. Then we have

$$h^{(k+1)}(t) = \left[\frac{1}{\hat{p_1}(1 + \hat{h_1}(t)t)^{-1} + \hat{p_2}(1 + \hat{h_2}t)^{-1}} - 1 \right] \frac{1}{t}.$$

Analogously to the case $n = 2$, this expression can be rearranged to give:

$$h^{(k+1)}(t) \;=\; \frac{\hat{p_1}\hat{h_1} + \hat{p_2}\hat{h_2} + \hat{h_1}\hat{h_2}t}{1 + \hat{p_1}\hat{h_2}t + \hat{p_2}\hat{h_1}t}.$$

However, by contrast to the case $n = 2$, $\hat{h_1}$ is now a function of t. The derivative of $h^{(k+1)}$ is:

$$\frac{dh^{(k+1)}(t)}{dt} =$$
$$\frac{\left(\hat{p_1}\hat{h}_1' + \hat{h_1}\hat{h_2} + \hat{h}_1'\hat{h_2}t\right)\left(1 + \hat{p_1}\hat{h_2}t + \hat{p_2}\hat{h_1}t\right) - \left(\hat{p_1}\hat{h_1} + \hat{p_2}\hat{h_2} + \hat{h_1}\hat{h_2}t\right)\left(\hat{p_1}\hat{h_2} + \hat{p_2}\hat{h_1} + \hat{p_2}\hat{h}_1't\right)}{\left[1 + \hat{p_1}\hat{h_2}t + \hat{p_2}\hat{h_1}t\right]^2}.$$

The denominator of this fraction is strictly positive, so the sign of the derivative depends on the numerator only. Denote the numerator by N:

$$\begin{aligned} N \;=\; & \left(\hat{p_1}\hat{h}_1' + \hat{h_1}\hat{h_2} + \hat{h}_1'\hat{h_2}t\right)\left(1 + \hat{p_1}\hat{h_2}t + \hat{p_2}\hat{h_1}t\right) \\ & - \left(\hat{p_1}\hat{h_1} + \hat{p_2}\hat{h_2} + \hat{h_1}\hat{h_2}t\right)\left(\hat{p_1}\hat{h_2} + \hat{p_2}\hat{h_1} + \hat{p_2}\hat{h}_1't\right). \end{aligned}$$

Note that

$$N \;=\; \hat{Q}(t) + \hat{h}_1'\left[\left(\hat{p}_1 + \hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \hat{p}_2t\left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right)\right],$$

where $\hat{Q}(t)$ is defined as in the proof of Proposition 1, but with $h_1 = \hat{h}_1(t)$ and $h_2 = \hat{h}_2$. Since Proposition 1 establishes that $\hat{Q}(t) \leq 0$ (with equality if and only if $\hat{h}(t) = h_2$) and $\hat{h}_1' < 0$, it suffices to show that

$$\left(\hat{p}_1 + \hat{h}_2t\right)\left(1 + \hat{p}_1\hat{h}_2t + \hat{p}_2\hat{h}_1t\right) - \hat{p}_2t\left(\hat{p}_1\hat{h}_1 + \hat{p}_2\hat{h}_2 + \hat{h}_1\hat{h}_2t\right) > 0 \qquad (3.21)$$

Cancelling terms on the left-hand side of (3.21) leaves us with:

$$\hat{p}_1\left(1 + \hat{p}_1\hat{h}_2t\right) + \hat{h}_2t\left(1 + \hat{p}_1\hat{h}_2t\right) - (\hat{p}_2)^2\,\hat{h}_2t.$$

We now use the fact that $(\hat{p}_2)^2 = (1 - \hat{p}_1)^2 = 1 - 2\hat{p}_1 + (\hat{p}_1)^2$ to get

$$\begin{aligned} & \hat{p}_1\left(1 + \hat{p}_1\hat{h}_2t\right) + \hat{h}_2t\left(1 + \hat{p}_1\hat{h}_2t\right) - \left[1 - 2\hat{p}_1 + (\hat{p}_1)^2\right]\hat{h}_2t \\ & = \; \hat{p}_1 + \left(t\hat{h}_2\right)^2\hat{p}_1 + 2\hat{p}_1\hat{h}_2t, \end{aligned}$$

which is strictly positive as required. Therefore, $h^{(k+1)}(t)$ is strictly decreasing. □

3.4 Discussion

We generalized (the continuous-time analogue of) Jackson and Yariv's result [46] by proving that whenever we aggregate different discount functions from from the same DI class, the weighted average function is always *strictly more DI* than each of its constituents. This also strengthens the conclusion of Corollary 4 in Prelec [63] who demonstrates that the mixture of two different discount functions from the same DI class represents more DI preferences.

We also show that mixing discount functions which are ordered by decreasing impatience produces a mixture that is more decreasingly impatient than the *least decreasingly impatient* component (Theorem 3.22). Surprisingly, in the light of Theorem 3.22, it need not be *strictly* so. This is confirmed by Example 3.21.

When a decision-maker is uncertain about her hyperbolic discount rate, we showed that long-term costs and benefits must be discounted at the probability-weighted harmonic mean of the possible hyperbolic discount rates. This complements the well-known result of Weitzman [81].

One natural question that arises is whether it is possible to prove a result analogous to Proposition 3.18 when all preference orders exhibit increasing impatience (II). The answer to this question can be partially found in the literature on survival analysis and reliability theory. In general, increasing failure rates are not preserved under mixtures. Indeed, Gurland and Sethuraman [39, 40] provide striking examples of mixtures of very quickly increasing failure rates that are eventually decreasing.

4

One-switch discount functions

In this chapter we analyse the one-switch property for preferences over sequences of dated outcomes introduced in [12]. This property concerns the effect of adding a common delay to two such sequences: it says that the preference ranking of the delayed sequences is either independent of the delay, or else there is a unique delay such that one ranking prevails for shorter delays and the opposite ranking for longer delays. For preferences that have a discounted utility (DU) representation, Bell [12] argues that sums of exponentials are the only discount functions consistent with the one-switch property. In this chapter we prove that discount functions of the linear times exponential form also satisfy the one-switch property. We also show that preferences which have a DU representation with a linear times exponential discount function exhibit strictly increasing impatience.

The chapter is organized as follows. We start by giving some preliminaries in Section 4.1. Section 4.2 is devoted to revising Bell's characterisation [12, Proposition 2] of the discount functions that exhibit the one-switch property. We first discuss an ambiguity in Bell's [12] definition of this property, and distinguish a standard (strong) version from an alternative "weak" one-switch property. We show the discount functions consistent with the (standard) one-switch property are those which have the sum of exponentials or the linear time exponential form. We also explore the relationship between the one-switch property restricted to preferences over dated outcomes (i.e., elements of $\mathcal{A}_1$) and the impatience properties of such preferences.

In Section 4.3 we study the weak one-switch property. In the context of expected utility preferences over lotteries, where the one-switch property refers to the effect of wealth level on the ranking of two lotteries, the results in [12] imply that the weak one-switch property is equivalent to the standard one. In the intertemporal context, we establish that the equivalence also holds if we endow X with a mixture set structure and work in an environment similar to of Anscombe and Aumann in [8]. Finally, Section 4.4 summarizes the results.

4.1 Preliminaries

Consider preferences over sequences of dated outcomes. Let us recall our basic notation from Section 1.1.1. We work in a continuous time environment throughout this Chapter. Points in time are elements of the set $T = [0, \infty)$, where the present time corresponds to $t = 0$. The set of outcomes is initially assumed to be the interval $X = [0, \infty)$, though we will re-define X to be an arbitrary mixture set in Section 4.3.

Let $\mathcal{A}_n = \{ \ (\mathbf{x}, \mathbf{t}) \in X^n \times T^n \mid t_1 < t_2 < \ldots < t_n \ \}$ be the set of sequences with n dated outcomes. Define the set of alternatives $\mathcal{A}$ as follows: $\mathcal{A} = \cup_{n=1}^{\infty} \mathcal{A}_n$. The set $\mathcal{A}$ consists of all sequences of finitely many dated outcomes. Elements of $\mathcal{A}_1 \subseteq \mathcal{A}$ are called dated outcomes.

Consider a preference order $\succcurlyeq$ on the set of alternatives $\mathcal{A}$.

We say that U is a *discounted utility (DU) representation* for $\succcurlyeq$, if U represents $\succcurlyeq$ and there exist (u, D), such that $u \colon X \to \mathbb{R}$ is a utility function (continuous, strictly increasing, $u(0) = 0$), $D \colon T \to (0, 1]$ is a discount function (strictly decreasing, $D(0) = 1$ and $\lim_{t \to \infty} D(t) = 0$) and

$$U(\mathbf{x}, \mathbf{t}) = \sum_{i=1}^{n} D(t_i) u(x_i)$$

for all n and every $(\mathbf{x}, \mathbf{t}) \in \mathcal{A}_n$. Necessary and sufficient conditions for an DU representation were provided by Harvey [43, Theorem 2.1].

We denote the set of positive integers $\{1, 2, 3, \ldots\}$ by $\mathbb{N}$, and the set of non-negative integers $\{0, 1, 2, 3, \ldots\}$ by $\mathbb{N}_0$, so $\mathbb{N} \subset \mathbb{N}_0$.

4.2 The one-switch property

4.2.1 One-switch discount functions

For any sequence of dated outcomes $(\mathbf{x}, \mathbf{t}) \in \mathcal{A}_n$ and any delay $\sigma > 0$, let

$$(\mathbf{x}, \mathbf{t} + \sigma) = (\mathbf{x}, (t_1 + \sigma, \ldots, t_n + \sigma))$$

denote the delayed sequence.

Definition 4.1 ([12]). *We say that the preferences $\succcurlyeq$ on $\mathcal{A}$ exhibit the* one-switch property *if for every pair $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ the ranking of $(\mathbf{x}, \mathbf{t} + \sigma)$ and $(\mathbf{y}, \mathbf{s} + \sigma)$ is either independent of σ, or there exists $\sigma^* \geq 0$, such that*

$$\begin{aligned} &(\mathbf{x}, \mathbf{t} + \sigma) \succ (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma < \sigma^*, \\ &(\mathbf{x}, \mathbf{t} + \sigma) \prec (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma > \sigma^* \end{aligned}$$

or

$$\begin{aligned} &(\mathbf{x}, \mathbf{t} + \sigma) \prec (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma < \sigma^*, \\ &(\mathbf{x}, \mathbf{t} + \sigma) \succ (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma > \sigma^*. \end{aligned}$$

It is worth mentioning that Bell's [12] original verbal definition of the one-switch property uses the ambiguous word "preferred", which does not specify whether the preference order is used in a strong or weak sense. Bell's Lemma 3 [12] implicitly suggests that weak preference is intended, but this Lemma also shows that either interpretation leads to the same restriction on EU preferences. Abbas and Bell [3] introduce a formal definition which is explicit about preference being strict. Therefore, we define the one-switch property using strict preference order in this section.

The one-switch property can be stated in a weaker form, as follows:

Definition 4.2 ([12]). *We say that the preferences $\succcurlyeq$ on $\mathcal{A}$ exhibit the* weak one-switch property *if for every pair $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ the ranking of $(\mathbf{x}, \mathbf{t} + \sigma)$ and $(\mathbf{y}, \mathbf{s} + \sigma)$ is either independent of σ, or there exists $\sigma^* \geq 0$, such that*

$$\begin{aligned} &(\mathbf{x}, \mathbf{t} + \sigma) \succcurlyeq (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma < \sigma^*, \\ &(\mathbf{x}, \mathbf{t} + \sigma) \preccurlyeq (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma > \sigma^* \end{aligned}$$

or

$$\begin{aligned} &(\mathbf{x}, \mathbf{t} + \sigma) \preccurlyeq (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma < \sigma^*, \\ &(\mathbf{x}, \mathbf{t} + \sigma) \succcurlyeq (\mathbf{y}, \mathbf{s} + \sigma), \ \textit{for any } \sigma > \sigma^*. \end{aligned}$$

In other words, there do not exist $(\mathbf{x},\mathbf{t}),(\mathbf{y},\mathbf{s}) \in \mathcal{A}$ and σ,ε with $0<\sigma<\varepsilon$ such that

$$\begin{aligned}(\mathbf{x},\mathbf{t}) &\succ (\mathbf{y},\mathbf{s}),\\ (\mathbf{x},\mathbf{t}+\sigma) &\prec (\mathbf{y},\mathbf{s}+\sigma),\\ (\mathbf{x},\mathbf{t}+\varepsilon) &\succ (\mathbf{y},\mathbf{s}+\varepsilon),\end{aligned}$$

or with all strict preferences reversed.

In the intertemporal context it is not known whether this alternative "weak" version is equivalent (given a DU representation) to Definition 4.1. This question will be investigated is Section 4.3, where we adapt Bell's Lemma 3 [12] to the temporal setting. We demonstrate that the one-switch property and the weak one-switch property are equivalent in an intertemporal version of the Anscombe and Aumann (AA) [8] environment similar to that investigated in Chapter 2.

If preferences $\succcurlyeq$ on $\mathcal{A}$ have a DU representation (u,D), then the one-switch property means that for any $(\mathbf{x},\mathbf{t}),(\mathbf{y},\mathbf{s}) \in \mathcal{A}$ the function

$$\Delta(\sigma)=\sum_{i=1}^{n} D(t_i+\sigma)u(x_i)-\sum_{j=1}^{m} D(s_j+\sigma)u(y_j)$$

either has constant sign or else there is some $\sigma^* \geq 0$ such that $\Delta(\sigma)=0$ if and only if $\sigma=\sigma^*$ and $\Delta(\sigma')\Delta(\sigma'')>0$ if and only if $\sigma'\neq\sigma^*$ and $\sigma''\neq\sigma^*$ are on the same side of σ^*. That is,

$$\begin{aligned}&\mathit{sign}\,(\Delta(\sigma))=\mathit{const} \text{ for all } \sigma\geq 0, \text{ or else}\\ &\text{there exists } \sigma^* \text{ such that } \Delta(\sigma^*)=0 \text{ and}\\ &\Delta(\sigma')\Delta(\sigma'')>0 \text{ if } \sigma',\sigma''>\sigma^* \text{ or } \sigma',\sigma''<\sigma^* \text{ and}\\ &\Delta(\sigma')\Delta(\sigma'')<0 \text{ if } \sigma'<\sigma^*<\sigma'' \text{ or } \sigma''<\sigma^*<\sigma'.\end{aligned} \tag{4.1}$$

Figure 4.1 provides an illustration of $\Delta(\sigma)$ for preferences which exhibit the one-switch property and have a DU representation. Note that only the sign of $\Delta(\sigma)$ is relevant.

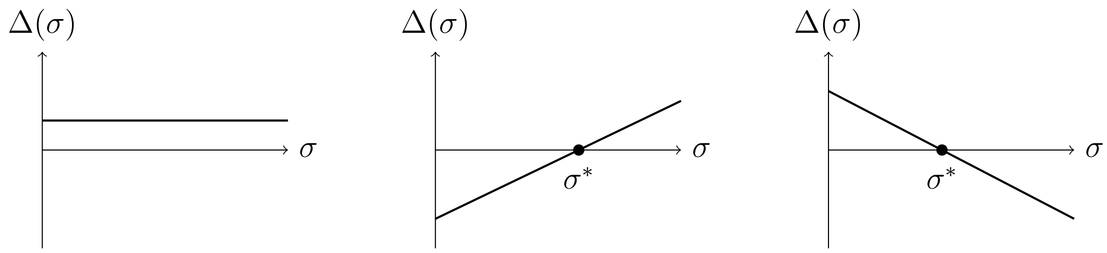

Figure 4.1: Possible behaviour of $\Delta(\sigma)$ for the preferences which exhibit the one-switch property and have a DU representation

Note that $u(X) = [0, \bar{u}]$ for some $\bar{u} > 0$ or $u(X) = \mathbb{R}_+$. We say that a discount function D satisfies the *extended* one-switch property if the function $\Delta\colon \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\Delta(\sigma) = \sum_{i=1}^{n} D(t_i + \sigma)v_i - \sum_{j=1}^{m} D(s_j + \sigma)w_j \tag{4.2}$$

satisfies (4.1) for any n, m, any $\mathbf{t} \in T^n, \mathbf{s} \in T^m$ and any $\mathbf{v} \in \mathbb{R}_+^n, \mathbf{w} \in \mathbb{R}_+^m$.

Lemma 4.3. *D satisfies the extended one-switch property if and only if there exists $\bar{u} > 0$ such that* (4.2) *satisfies* (4.1) *for any n, m, any $\mathbf{t} \in T^n, \mathbf{s} \in T^m$ and any $\mathbf{v} \in [0, \bar{u}]^n, \mathbf{w} \in [0, \bar{u}]^m$.*

Proof. "Only If". This part is straightforward.

"If". Suppose that there exists $\bar{u} > 0$ such that $\Delta : \mathbb{R}_+ \to \mathbb{R}$ defined by (4.2) satisfies (4.1) for any n, m, any $\mathbf{t} \in T^n, \mathbf{s} \in T^m$ and any $\mathbf{v} \in [0, \bar{u}]^n, \mathbf{w} \in [0, \bar{u}]^m$. The proof is by contradiction. Assume that there is some n', m', $\mathbf{t}' \in T^{n'}, \mathbf{s}' \in T^{m'}$ and some $\mathbf{v}' \in \mathbb{R}_+^{n'}, \mathbf{w}' \in \mathbb{R}_+^{m'}$ such that the function $\Delta^*\colon \mathbb{R}_+ \to \mathbb{R}$ defined by

$$\Delta^*(\sigma) = \sum_{i=1}^{n'} D(t'_i + \sigma)v'_i - \sum_{j=1}^{m'} D(s'_j + \sigma)w'_j.$$

violates property (4.1). Then the function $\lambda\Delta^*$ will also violate (4.1) for any $\lambda > 0$. Let $\lambda \in (0, 1)$ be such that $\lambda\mathbf{v}' \in [0, \bar{u}]^{n'}$ and $\lambda\mathbf{w}' \in [0, \bar{u}]^{m'}$. This is a contradiction to the initial assumption that (4.2) satisfies (4.1) for any n, m, any $\mathbf{t} \in T^n, \mathbf{s} \in T^m$ and any $\mathbf{v} \in [0, \bar{u}]^n, \mathbf{w} \in [0, \bar{u}]^m$. □

In other words, Lemma 4.3 states that given preferences with a DU representation, the range of u is irrelevant to whether or not the preferences satisfy the one-switch property. It follows that the one-switch property does not impose any additional restrictions on the shape of u. In other words, given a DU representation (u, D) for $\succcurlyeq$, the one-switch property restricts only D. We therefore say that a discount function, D, exhibits the one-switch property if there is some utility function, u, such that the preferences with DU representation (u, D) exhibit the one-switch property. Bell [12, Proposition 8] argues that sums of exponentials are the only discount functions compatible with the one-switch property. However, we will demonstrate in this section that linear times exponential discount functions also have the one-switch property.

The following proposition gives the restrictions on the parameters of linear times exponential and sum of exponential functions under which they satisfy the properties of a discount function.

Proposition 4.4. *(a) The linear times exponential function* $D(t) = (c_1 + c_2 t)e^{r_1 t}$ *is a discount function if and only if* $c_1 = 1, -r_1 \geq c_2 \geq 0$ *and* $r_1 < 0$*; i.e.,* $D(t) = (1 + ct)e^{-rt}$*, where* $r \geq c \geq 0$ *and* $r > 0$.

(b) The sum of exponentials function $D(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$*, where* $r_1 \neq r_2$*, is a discount function if and only if*

- $c_2 = 1 - c_1, r_1 < r_2 < 0$ *and* $\frac{r_2}{r_2 - r_1} \leq c_1 < 1$*, or*
- $c_2 = 1 - c_1, r_2 < r_1 < 0$ *and* $0 < c_1 \leq \frac{r_2}{r_2 - r_1}$.

Equivalently, $D(t) = ae^{-bt} + (1 - a)e^{-(b+c)t}$*, where* $a, b, c > 0$*,* $a \leq b/c + 1$.

Proof. We need to find the parameters of linear times exponential functions and sums of exponentials such that the following four properties are satisfied:

1. $D(0) = 1$,
2. $D(t) > 0$ for all t,
3. $D(t)$ is strictly decreasing, and
4. $\lim_{t \to \infty} D(t) = 0$.

(a) Linear times exponential. Assume that $D(t) = (c_1 + c_2 t)e^{r_1 t}$. Obviously, $D(0) = 1$ if and only if $c_1 = 1$.

Next, to satisfy the first and the second conditions simultaneously, that is, to have $D(0) = 1$ and $D(t) > 0$ for all t, it is necessary and sufficient that $c_1 = 1$ and $c_2 \geq 0$ (since $e^{r_1 t} > 0$ for all t).

To check whether $D(t)$ is strictly decreasing, consider its first order derivative:

$$D'(t) \;=\; c_2 e^{r_1 t} + (1 + c_2 t) r_1 e^{r_1 t} \;=\; e^{r_1 t}\left(c_2 + r_1 + c_2 r_1 t\right).$$

Since $e^{rt} > 0$ for all t, the sign of the derivative depends on the sign of the linear expression $c_2 + r_1 + c_2 r_1 t$. Therefore, $D(t)$ is strictly decreasing if and only if

$$c_2 + r_1 \leq 0, \text{ and}$$
$$c_2 + r_1 + c_2 r_1 t < 0 \text{ for all } t > 0.$$

This condition is equivalent to requiring that one of the following two conditions is satisfied:

- $c_2 + r_1 = 0$ and $c_2 r_1 < 0$,
- $c_2 + r_1 < 0$ and $c_2 r_1 \leq 0$.

Since $c_2 \geq 0$, in the first case we have $c_2 = -r_1 > 0$. In the second case $0 \leq c_2 < -r_1$. We can summarize both cases as follows:

$$-r_1 \geq c_2 \geq 0 \text{ and } r_1 < 0.$$

Therefore, the first three properties of discount functions are satisfied if and only if $c_1 = 1, -r_1 \geq c_2 \geq 0$ and $r_1 < 0$.

Finally, the limit of the linear times exponential function is

$$\lim_{t\to\infty} D(t) \;=\; \lim_{t\to\infty} (1 + c_2 t)\, e^{r_1 t} \;=\; \lim_{t\to\infty} \frac{1 + c_2 t}{e^{-r_1 t}},$$

If $r_1 = c_2 = 0$ this limit is 1. This case is ruled out since from the previous step $r_1 < 0$. Then, by L'Hopital's rule we have

$$\lim_{t\to\infty} D(t) \;=\; \lim_{t\to\infty} \frac{1 + c_2 t}{e^{-r_1 t}} \;=\; \lim_{t\to\infty} \frac{c_2}{-r_1 e^{-r_1 t}} \;=\; 0.$$

Therefore, $D(t) = (c_1 + c_2 t)e^{r_1 t}$ satisfies all four properties of discount functions if and only if $c_1 = 1, -r_1 \geq c_2 \geq 0$ and $r_1 < 0$. Denote $c = c_2$ and $r = -r_1$. Then we have $D(t) = (1 + ct)e^{-rt}$, where $r \geq c \geq 0$ and $r > 0$.

(b) Sums of exponentials. The proof is analogous to [12, Proposition 8] and is given here for completeness. Assume that $D(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$ with $r_1 \neq r_2$.

The condition $D(0) = 1$ is satisfied if and only if $c_1 + c_2 = 1$.

We must also have $D(t) > 0$ for all $t > 0$. Note that

$$D(t) \;=\; c_1 e^{r_1 t} + c_2 e^{r_2 t} \;=\; e^{r_1 t}\left(c_1 + c_2 e^{(r_2 - r_1)t}\right) > 0 \text{ for all } t > 0$$

if and only if $c_1 + c_2 e^{(r_2 - r_1)t} > 0$ for all $t > 0$ (since $e^{r_1 t} > 0$ for all $t > 0$). From the first condition we know that $c_2 = 1 - c_1$. Substituting this expression to the inequality we must have $c_1 + (1 - c_1)e^{(r_2 - r_1)t} > 0$ for all $t > 0$. Therefore, the first two properties of discount functions are satisfied if and only if $c_2 = 1 - c_1$ and one of the following two conditions holds:

(i) $r_1 < r_2$ and $c_1 \leq 1$; or

(ii) $r_1 > r_2$ and $c_1 \geq 0$.

Next, it is necessary to have $D(t)$ strictly decreasing. Consider its first order derivative

$$D'(t) \;=\; c_1 r_1 e^{r_1 t} + c_2 r_2 e^{r_2 t} \;=\; e^{r_1 t}\left(c_1 r_1 + c_2 r_2 e^{(r_2 - r_1)t}\right).$$

Since $e^{r_1 t} > 0$ for all $t > 0$, the function $D(t)$ is strictly decreasing if and only if

$$\begin{gathered} c_1 r_1 + c_2 r_2 \leq 0, \text{ and} \\ c_1 r_1 + c_2 r_2 e^{(r_2 - r_1)t} < 0 \text{ for all } t > 0. \end{gathered}$$

Recall that from the first two conditions we have $c_2 = 1 - c_1$ and [(i) or (ii)] holds. Therefore, $D(0) = 1, D(t) > 0$ for all $t > 0$ and $D(t)$ is strictly decreasing if and only if all of the following conditions hold:

$$\begin{aligned} & c_1 r_1 + (1 - c_1) r_2 \leq 0, \\ & c_1 r_1 + (1 - c_1) r_2 e^{(r_2 - r_1)t} < 0 \text{ for all } t > 0, \\ & c_2 = 1 - c_1, \text{ and [(i) or (ii)] holds.} \end{aligned}$$

Note that

$$\begin{aligned} c_1 r_1 + (1 - c_1) r_2 e^{(r_2 - r_1)t} &= c_1 r_1 + (1 - c_1) r_2 e^{(r_2 - r_1)t} + c_1 r_1 e^{(r_2 - r_1)t} - c_1 r_1 e^{(r_2 - r_1)t} \\ &= \left(c_1 r_1 + (1 - c_1) r_2\right) e^{(r_2 - r_1)t} + c_1 r_1 \left(1 - e^{(r_2 - r_1)t}\right). \end{aligned}$$

Therefore, we must have

$$c_1 r_1 + (1 - c_1) r_2 \leq 0, \tag{4.3}$$

$$\left(c_1 r_1 + (1 - c_1) r_2\right) e^{(r_2 - r_1)t} + c_1 r_1 \left(1 - e^{(r_2 - r_1)t}\right) < 0 \text{ for all } t > 0, \tag{4.4}$$

$$c_2 = 1 - c_1, \text{ and [(i) or (ii)] holds.} \tag{4.5}$$

Consider case (i) of (4.5). Then condition (4.3) holds if and only if $c_1 \geq \frac{r_2}{r_2 - r_1}$. Given (4.3), condition (4.4) holds if and only if $(1 - c_1) r_2 < 0$, which is equivalent – in case (i) – to $c_1 < 1$ and $r_2 < 0$. Thus, in case (i), the first three properties of a discount function are satisfied if and only if $c_2 = 1 - c_1, r_1 < r_2 < 0$ and $\frac{r_2}{r_2 - r_1} \leq c_1 < 1$. Since $r_1 < r_2 < 0$, the fourth property is also satisfied.

Next, consider case (ii) of (4.5). The argument is similar to case (i). The condition (4.3) holds if and only if $c_1 \leq \frac{r_2}{r_2 - r_1}$. Given (4.3), condition (4.4) holds if and only if $c_1 r_1 < 0$, which is equivalent – in case (ii) – to $c_1 > 0$ and $r_1 < 0$. Therefore, in case (ii), the first three properties of a discount function are satisfied if and only if $c_2 = 1 - c_1, r_2 < r_1 < 0$ and $0 < c_1 \leq \frac{r_2}{r_2 - r_1}$. Since $r_2 < r_1 < 0$, the fourth property is also satisfied.

In case (i) of (4.5) let $a = c_2, b = -r_2$, and $c = r_2 - r_1$. It follows that $a, b, c > 0$. Since $r_2 / r_2 - r_1 \leq c_1 < 1$ it requires that $0 < a \leq b/c + 1$. We then have $D(t) = a e^{-bt} + (1 - a) e^{-(b+c)t}$, where $a, b, c > 0$, $a \leq b/c + 1$.

Analogously, in case (ii) of (4.5) let $a = c_1, b = -r_1$, and $c = r_1 - r_2$. We then obtain the same functional from and parameter restrictions as in case (i), that is, $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, where $a, b, c > 0$, $a \leq b/c + 1$. □

The following proposition demonstrates that these two types of discount function are compatible with the one-switch property.

Proposition 4.5. *Suppose that the preference order $\succcurlyeq$ on $\mathcal{A}$ has a DU representation (u, D), where*

- $D(t) = (1 + ct)e^{-rt}$, *where $r \geq c \geq 0$ and $r > 0$, or*
- $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, *where $a, b, c > 0$, and $a \leq b/c + 1$.*

Then $\succcurlyeq$ exhibits the one-switch property.

Proof. The proof adapts Bell's argument [12, Proposition 2] to the time preference framework. We need to prove that for any $(\mathbf{x}, \mathbf{t})$, $(\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ the following function changes sign at most once:

$$\Delta(\sigma) = \sum_{i=1}^{n} D(t_i + \sigma)u(x_i) - \sum_{j=1}^{m} D(s_j + \sigma)u(y_j).$$

(a) Linear times exponential. Consider the discount function $D(t) = (1 + ct)e^{-rt}$, where $r \geq c \geq 0$ and $r > 0$. Then

$$\Delta(\sigma) \;=\; \sum_{i=1}^{n}(1 + ct_i + c\sigma)e^{-rt_i}e^{-r\sigma}u(x_i) - \sum_{j=1}^{m}(1 + cs_j + c\sigma)e^{-rs_j}e^{-r\sigma}u(y_j).$$

Rearranging,

$$\begin{aligned}
\Delta(\sigma) \;&= e^{-r\sigma}\left(\sum_{i=1}^{n} e^{-rt_i}u(x_i) - \sum_{j=1}^{m} e^{-rs_j}u(y_j)\right) \\
&\quad + e^{-r\sigma}c\left(\sum_{i=1}^{n} t_i e^{-rt_i}u(x_i) - \sum_{j=1}^{m} s_j e^{-rs_j}u(y_j)\right) \\
&\quad + e^{-r\sigma}c\sigma\left(\sum_{i=1}^{n} e^{-rt_i}u(x_i) - \sum_{j=1}^{m} e^{-rs_j}u(y_j)\right).
\end{aligned}$$

Let

$$A \;=\; \sum_{i=1}^{n} e^{-rt_i}u(x_i) - \sum_{j=1}^{m} e^{-rs_j}u(y_j),$$

and

$$B = \sum_{i=1}^{n} t_i e^{-rt_i} u(x_i) - \sum_{j=1}^{m} s_j e^{-rs_j} u(y_j).$$

Then

$$\Delta(\sigma) = Ae^{-r\sigma} + cBe^{-r\sigma} + cAe^{-r\sigma}\sigma.$$

This expression can be rewritten as follows

$$\Delta(\sigma) = e^{-r\sigma}\left(A + cB + cA\sigma\right).$$

Since $e^{-r\sigma} > 0$, the sign of $\Delta(\sigma)$ equals the sign of $A + cB + cA\sigma$. Since the latter is linear its sign is constant or else changes once at a unique σ value.

(b) Sums of exponentials. Consider the function $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, where $a, b, c > 0$, and $a \le b/c + 1$. Then

$$\begin{aligned}\Delta(\sigma) = & \sum_{i=1}^{n}\left(ae^{-bt_i}e^{-b\sigma} + (1-a)e^{-(b+c)t_i}e^{-(b+c)\sigma}\right)u(x_i) \\ & - \sum_{j=1}^{m}\left(ae^{-bs_j}e^{-b\sigma} + (1-a)e^{-(b+c)s_j}e^{-(b+c)\sigma}\right)u(y_j).\end{aligned}$$

It can be rearranged so that

$$\begin{aligned}\Delta(\sigma) = & \; ae^{-b\sigma}\left(\sum_{i=1}^{n} e^{-bt_i}u(x_i) - \sum_{j=1}^{m} e^{-bs_j}u(y_j)\right) \\ & + (1-a)e^{-(b+c)\sigma}\left(\sum_{i=1}^{n} e^{-(b+c)t_i}u(x_i) - \sum_{j=1}^{m} e^{-(b+c)s_j}u(y_j)\right).\end{aligned}$$

Denote

$$\tilde{A} = \sum_{i=1}^{n} e^{-bt_i}u(x_i) - \sum_{j=1}^{m} e^{-bs_j}u(y_j),$$

and

$$\tilde{B} = \sum_{i=1}^{n} e^{-(b+c)t_i}u(x_i) - \sum_{j=1}^{m} e^{-(b+c)s_j}u(y_j).$$

Then

$$\Delta(\sigma) = a\tilde{A}e^{-b\sigma} + (1-a)\tilde{B}e^{-(b+c)\sigma}.$$

This expression can be factorized as follows

$$\Delta(\sigma) = e^{-b\sigma}\left(a\tilde{A} + (1-a)\tilde{B}e^{-c\sigma}\right).$$

Since $e^{-b\sigma} > 0$, the sign of $\Delta(\sigma)$ equals the sign of $a\tilde{A} + (1-a)\tilde{B}e^{-c\sigma}$. Therefore, $\Delta(\sigma)$ is either constant or else changes once at unique σ value. □

Since Proposition 4.5 establishes that linear times exponential discount functions and sum of exponentials discount functions satisfy the one-switch property, they must also satisfy the weak one-switch property.

4.2.2 The one-switch property for dated outcomes and monotonic impatience

In this section we consider preferences $\succcurlyeq$ on the set $\mathcal{A}_1$ of dated outcomes. When the preferences $\succcurlyeq$ are restricted to $\mathcal{A}_1$, then a DU representation becomes $U(x,t) = D(t)u(x)$ for any $(x,t) \in \mathcal{A}_1$. Necessary and sufficient conditions for this representation are given in [32]. We assume that $\succcurlyeq$ satisfy Fishburn and Rubinstein's axioms [32] throughout this section. The reader will find these axioms listed in Section 1.1.1 of Chapter 1.

Definition 4.6. *We say that $\succcurlyeq$ exhibits the* one-switch property for dated outcomes *if $\succcurlyeq$ exhibits the one-switch property on $\mathcal{A}_1$.*[1]

Obviously, if $\succcurlyeq$ exhibits the one-switch property on $\mathcal{A}$, it implies that $\succcurlyeq$ also exhibits the one-switch property for dated outcomes.

Recall the following notions of decreasing and increasing impatience.

Definition 4.7 ([63])**.** *We say that $\succcurlyeq$ exhibits* [strictly] decreasing impatience, *if for all $(x,t), (y,s) \in \mathcal{A}_1$ such that $0 < x < y$ and for all $t < s$: if $(x,t) \sim (y,s)$ then for any $\sigma > 0$ we have*

$$(x, t+\sigma)\ [\prec]\ \preccurlyeq (y, s+\sigma). \tag{4.6}$$

We say that $\succcurlyeq$ exhibits

- [strictly] increasing impatience*, if the preference in* (4.6) *is reversed;*
- stationarity, or constant impatience*, if the preference in* (4.6) *is replaced by indifference.*

When preferences have a DU representation, these properties only restrict the discount function. The next proposition follows directly from the definition.

Proposition 4.8. *Suppose that $\succcurlyeq$ restricted to $\mathcal{A}_1$ has a DU representation. Then $\succcurlyeq$ exhibits [strictly] DI if and only if*

[1] The related concept of an ordinal one-switch utility function is introduced in [3].

$$\frac{D(t)}{D(t+\sigma)} \; [>] \geq \frac{D(s)}{D(s+\sigma)}, \; \textit{for all } t, s \textit{ such that } t < s, \textit{ and every } \sigma > 0. \qquad (4.7)$$

Furthermore, $\succcurlyeq$ *exhibits*

- *[strictly] II if and only if the inequality in* (4.7) *is reversed;*
- *constant impatience if and only if the inequality in* (4.7) *is replaced by the equality.*

The following proposition provides a complete characterisation.

Proposition 4.9 ([63], [7]). *Suppose that* $\succcurlyeq$ *restricted to* $\mathcal{A}_1$ *has a DU representation. Then* $\succcurlyeq$ *exhibits*

- *[strictly] DI if and only if* $D(t)$ *is [strictly] log-convex,*[2]
- *[strictly] II if and only if* $D(t)$ *is [strictly] log-concave,*
- *constant impatience if and only if* $D(t) = e^{-rt}$ *with* $r > 0$.

Proposition 4.9 extends [63, Corollary 1] to increasing impatience and strictly decreasing (and strictly increasing) impatience. The proof is omitted here, since it only requires a minor adjustment of Prelec's original proof.[3]

The following lemma re-expresses the definition of the one-switch property for dated outcomes in terms of a common advancement ($\sigma < 0$) and a common delay ($\sigma > 0$) applied to a pair of dated outcomes between which the decision-maker is indifferent.

Lemma 4.10. *Let* $(\hat{x}, \hat{t}), (\hat{y}, \hat{s}) \in \mathcal{A}_1$ *such that* $0 < \hat{x} < \hat{y}$, $0 < \hat{t} < \hat{s}$ *and* $(\hat{x}, \hat{t}) \sim (\hat{y}, \hat{s})$. *Then* $\succcurlyeq$ *exhibits the one-switch property for dated outcomes only if either*

(i) $(\hat{x}, \hat{t}+\sigma) \sim (\hat{y}, \hat{s}+\sigma)$ *for all* $\sigma \geq -\hat{t}$, *or*

(ii) $(\hat{x}, \hat{t}+\sigma) \succ (\hat{y}, \hat{s}+\sigma)$ *for* $-\hat{t} \leq \sigma < 0$, *and* $(\hat{x}, \hat{t}+\sigma) \prec (\hat{y}, \hat{s}+\sigma)$ *for* $\sigma > 0$, *or*

(iii) $(\hat{x}, \hat{t}+\sigma) \prec (\hat{y}, \hat{s}+\sigma)$ *for* $-\hat{t} \leq \sigma < 0$, *and* $(\hat{x}, \hat{t}+\sigma) \succ (\hat{y}, \hat{s}+\sigma)$ *for* $\sigma > 0$.

Proof. Let $(\hat{x}, \hat{t}), (\hat{y}, \hat{s}) \in \mathcal{A}_1$ such that $0 < \hat{x} < \hat{y}$, $0 < \hat{t} < \hat{s}$ and $(\hat{x}, \hat{t}) \sim (\hat{y}, \hat{s})$. Assume also that $\succcurlyeq$ exhibits the one-switch property for dated outcomes. Therefore, by definition of the one-switch property we have

(i*) $(\hat{x}, \hat{t}+\sigma) \sim (\hat{y}, \hat{s}+\sigma)$ for all $\sigma > 0$, or

(ii*) $(\hat{x}, \hat{t}+\sigma) \prec (\hat{y}, \hat{s}+\sigma)$ for all $\sigma > 0$, or

---

[2] A function $f\colon I \to \mathbb{R}$ is called *log-convex* if $f(x) > 0$ for all $x \in I$ and $\ln(f)$ is convex; and *strictly log-convex* if $f(x) > 0$ for all $x \in I$ and $\ln(f)$ is strictly convex. We say that a function $f\colon I \to \mathbb{R}$ is *[strictly] log-concave*, if $1/f$ is [strictly] log-convex. See Section 3.1.1 of Chapter 3.

[3] The proof of the proposition can also be found in the working paper [7]. See Proposition 3.5 in Section 3.1.2 of Chapter 3.

(iii*) $(\hat{x}, \hat{t}+\sigma) \succ (\hat{y}, \hat{s}+\sigma)$ for all $\sigma > 0$.

We need to analyse the situation when $-\hat{t} \leq \sigma < 0$. The proof is by contradiction.

In case (i*), assume that there exists μ^* such that $0 < \mu^* < \hat{t}$ and $(\hat{x}, \hat{t}-\mu^*) \prec (\hat{y}, \hat{s}-\mu^*)$. Let $\hat{\tau} = \hat{t}-\mu^* > 0$ and $\hat{\rho} = \hat{s}-\mu^* > 0$. Using this notation, we obtain

$$(\hat{x}, \hat{\tau}) \prec (\hat{y}, \hat{\rho}),$$
$$(\hat{x}, \hat{\tau}+\mu^*+\sigma) \sim (\hat{y}, \hat{\rho}+\mu^*+\sigma), \text{ for all } \sigma \geq 0.$$

This contradicts the one-switch property for dated outcomes, so (i) follows. If $(\hat{x}, \hat{t}-\mu^*) \succ (\hat{y}, \hat{s}-\mu^*)$ the proof is analogous.

In case (ii*), assume that there exists μ^* such that $0 < \mu^* < \hat{t}$ and $(\hat{x}, \hat{t}-\mu^*) \preccurlyeq (\hat{y}, \hat{s}-\mu^*)$. With the same notation as in the previous case $\hat{\tau} = \hat{t}-\mu^* > 0$ and $\hat{\rho} = \hat{s}-\mu^* > 0$ gives us

$$(\hat{x}, \hat{\tau}) \preccurlyeq (\hat{y}, \hat{\rho}),$$
$$(\hat{x}, \hat{\tau}+\mu^*) \sim (\hat{y}, \hat{\rho}+\mu^*),$$
$$(\hat{x}, \hat{\tau}+\mu^*+\sigma) \prec (\hat{y}, \hat{\rho}+\mu^*+\sigma), \text{ for all } \sigma > 0,$$

which is a contradiction.

In case (iii*) the proof is symmetric to case (ii*). □

The relation between impatience properties and the one-switch property for dated outcomes is established in the following lemma.

Lemma 4.11. *Suppose that $\succcurlyeq$ has a DU representation (u, D). Then $\succcurlyeq$ exhibits the one-switch property for dated outcomes if and only if $\succcurlyeq$ exhibits either stationarity or strictly DI or strictly II.*

Proof. "Only If". Assume that $\succcurlyeq$ exhibits the one-switch property for dated outcomes.

Consider some $(\hat{x}, \hat{t}), (\hat{y}, \hat{s}) \in \mathcal{A}_1$ such that $0 < \hat{x} < \hat{y}$, $0 < \hat{t} < \hat{s}$ and $(\hat{x}, \hat{t}) \sim (\hat{y}, \hat{s})$. To see that we can always find such a pair, suppose that $0 < \hat{x} < \hat{y}$, $\hat{t} < \hat{s}$ and $(\hat{x}, \hat{t}) \succ (\hat{y}, \hat{s})$. Then it follows by the continuity of u and the fact that D is strictly decreasing that there exists $t' \in (\hat{t}, \hat{s})$ such that $(\hat{x}, t') \sim (\hat{y}, \hat{s})$. Alternatively, suppose that $0 < \hat{x} < \hat{y}$, $\hat{t} < \hat{s}$ and $(\hat{x}, \hat{t}) \prec (\hat{y}, \hat{s})$. Then it follows by the continuity of u and the fact that D is strictly decreasing that there exists $x' \in (\hat{x}, \hat{y})$ such that $(x', \hat{t}) \sim (\hat{y}, \hat{s})$.

It follows by the one-switch property for dated outcomes and Lemma 4.10 that either

Case 1. $(\hat{x}, \hat{t}+\sigma) \sim (\hat{y}, \hat{s}+\sigma)$ for all $\sigma \geq -\hat{t}$, or

Case 2. $(\hat{x}, \hat{t}+\sigma) \succ (\hat{y}, \hat{s}+\sigma)$ for $-\hat{t} \leq \sigma < 0$, and $(\hat{x}, \hat{t}+\sigma) \prec (\hat{y}, \hat{s}+\sigma)$ for $\sigma > 0$, or

Case 3. $(\hat{x}, \hat{t}+\sigma) \prec (\hat{y}, \hat{s}+\sigma)$ for $-\hat{t} \leq \sigma < 0$, and $(\hat{x}, \hat{t}+\sigma) \succ (\hat{y}, \hat{s}+\sigma)$ for $\sigma > 0$.

We will analyse each case separately.

Case 1. Note that letting $\alpha = \sigma + \hat{t} \geq 0$ and $\hat{\sigma} = \hat{s} - \hat{t} > 0$ we have

$$(\hat{x}, \alpha) \sim (\hat{y}, \hat{\sigma} + \alpha) \text{ for all } \alpha \geq 0.$$

Using the DU representation it follows that

$$\frac{u(\hat{x})}{u(\hat{y})} = \frac{D(\alpha + \hat{\sigma})}{D(\alpha)} \text{ for all } \alpha \geq 0. \tag{4.8}$$

Consider some $t' < s'$. Then (4.8) implies

$$\frac{u(\hat{x})}{u(\hat{y})} = \frac{D(t' + \hat{\sigma})}{D(t')} = \frac{D(s' + \hat{\sigma})}{D(s')}.$$

Rearranging,

$$\frac{D(s')}{D(t')} = \frac{D(s' + \hat{\sigma})}{D(t' + \hat{\sigma})}. \tag{4.9}$$

By continuity we can choose $x' < y'$ such that

$$\frac{D(s')}{D(t')} = \frac{u(x')}{u(y')}. \tag{4.10}$$

Hence, it follows from (4.9), (4.10) that

$$(x', t') \sim (y', s') \text{ and } (x', t' + \hat{\sigma}) \sim (y', s' + \hat{\sigma}).$$

Then the one-switch property implies that

$$\frac{D(s')}{D(t')} = \frac{D(s' + \mu)}{D(t' + \mu)} \text{ for all } \mu > 0.$$

Since $t' < s'$ were arbitrary, it follows by Proposition 4.8 that $\succcurlyeq$ exhibits constant impatience.

Case 2. Defining $\alpha, \hat{\sigma}$ as for Case 1, we have

$$(\hat{x}, \alpha) \succ (\hat{y}, \hat{\sigma} + \alpha) \text{ for } 0 \leq \alpha < \hat{t}, \text{ and}$$
$$(\hat{x}, \alpha) \prec (\hat{y}, \hat{\sigma} + \alpha) \text{ for } \sigma > \hat{t}.$$

Therefore, using the DU representation

$$\frac{u(\hat{x})}{u(\hat{y})} > \frac{D(\hat{\sigma} + \alpha)}{D(\alpha)} \text{ for } 0 \leq \alpha < \hat{t}, \text{ and}$$

$$\frac{u(\hat{x})}{u(\hat{y})} < \frac{D(\hat{\sigma}+\alpha)}{D(\alpha)} \text{ for } \alpha > \hat{t}.$$

Hence,

$$\frac{D(t'+\hat{\sigma})}{D(t')} < \frac{u(\hat{x})}{u(\hat{y})} < \frac{D(s'+\hat{\sigma})}{D(s')} \text{ for any } t' < \hat{t} < s'.$$

Rearranging

$$\frac{D(s')}{D(t')} < \frac{D(s'+\hat{\sigma})}{D(t'+\hat{\sigma})} \text{ for any } t' < \hat{t} < s'. \tag{4.11}$$

By continuity we can choose $x' < y'$ such that

$$\frac{D(s')}{D(t')} = \frac{u(x')}{u(y')}. \tag{4.12}$$

It follows from (4.11), (4.12) and the one-switch property that

$$\frac{D(s'-\mu)}{D(t'-\mu)} < \frac{D(s')}{D(t')} < \frac{D(s'+\mu)}{D(t'+\mu)} \tag{4.13}$$

for any $\mu > 0$ whenever $t' < \hat{t} < s'$.

Consider some $t' < s'$. There are three possible sub-cases:

(a) $t' < \hat{t} < s'$,

(b) $t' < s' \leq \hat{t}$, and

(c) $\hat{t} \leq t' < s'$.

We will show that, in each of these three sub-cases,

$$\frac{D(s')}{D(t')} < \frac{D(s'+\mu)}{D(t'+\mu)} \text{ for all } \mu > 0. \tag{4.14}$$

From Proposition 4.8 we may then conclude that the preferences exhibit strict DI.

In sub-case (a), it follows directly from (4.13) that (4.14) holds.

In sub-case (b) choose $\varepsilon > 0$ such that $s' + \varepsilon < \hat{t} < t' + \varepsilon$. Let $t'' = t' + \varepsilon$ and $s'' = s' + \varepsilon$. It follows from (4.13) that

$$\frac{D(s''-\sigma)}{D(t''-\sigma)} < \frac{D(s'')}{D(t'')} < \frac{D(s''+\sigma)}{D(t''+\sigma)} \tag{4.15}$$

for all $\sigma > 0$. Let $\sigma = \varepsilon$. Then we have

$$\frac{D(s')}{D(t')} < \frac{D(s'+\varepsilon)}{D(t'+\varepsilon)}. \tag{4.16}$$

By continuity we can choose $x' < y'$ such that

$$\frac{D(s')}{D(t')} = \frac{u(x')}{u(y')}. \tag{4.17}$$

Therefore, it follows from (4.16), (4.17) and the one-switch property that (4.14) holds.

In sub-case (c) choose $\varepsilon > 0$ such that $t' - \varepsilon < \hat{t} < s' - \varepsilon$. Let $t'' = t' - \varepsilon$ and $s'' = s' - \varepsilon$. Then it follows from (4.13) that

$$\frac{D(s'' - \sigma)}{D(t'' - \sigma)} < \frac{D(s'')}{D(t'')} < \frac{D(s'' + \sigma)}{D(t'' + \sigma)} \tag{4.18}$$

for all $\sigma > 0$. Let $\sigma = \varepsilon$. Then we have

$$\frac{D(s' - \varepsilon)}{D(t' - \varepsilon)} < \frac{D(s')}{D(t')}. \tag{4.19}$$

Choose $x' < y'$ such that

$$\frac{D(s')}{D(t')} = \frac{u(x')}{u(y')}. \tag{4.20}$$

Therefore, (4.14) follows by (4.19), (4.20) and the one-switch property.

Therefore, in Case 2 $\succcurlyeq$ exhibits strictly DI.

Finally, in *Case 3* we may show that $\succcurlyeq$ exhibits strictly II by a symmetric proof to that for Case 2.

"If". Suppose that there are some x, y, t, s with $t \leq s$ and some $\sigma^* \geq 0$ such that $(x, t + \sigma^*) \sim (y, s + \sigma^*)$. It suffices to show that either

$$(x, t + \sigma) \sim (y, s + \sigma) \text{ for all } \sigma, \tag{4.21}$$

or

$$(x, t + \sigma) \succ (y, s + \sigma) \text{ for all } \sigma < \sigma^*, \text{ and} \tag{4.22}$$

$$(x, t + \sigma) \prec (y, s + \sigma) \text{ for all } \sigma > \sigma^*. \tag{4.23}$$

If $t = s$ then $(x, t + \sigma^*) \sim (y, t + \sigma^*)$. It follows by monotonicity that $x = y$. Hence, $\succcurlyeq$ satisfies the one-switch property for dated outcomes.

Assume that $t < s$ and $\succcurlyeq$ exhibits constant impatience. It follows that $(x, t + \sigma^* + \sigma') \sim (y, s + \sigma^* + \sigma')$ for all $\sigma' > 0$, or $(x, t + \sigma) \sim (y, s + \sigma)$ for all $\sigma > \sigma^*$. To show that $\succcurlyeq$ satisfies the one-switch property for dated outcomes we need to demonstrate that $(x, t + \sigma) \sim (y, s + \sigma)$ for all $\sigma < \sigma^*$ such that $t + \sigma \geq 0$, or $(x, t + \sigma^* - \sigma') \sim (y, s + \sigma^* - \sigma')$ for all $\sigma' > \sigma^*$ such that $t + \sigma^* - \sigma' \geq 0$. The proof is by contradiction. Suppose that there exists $\sigma'' > \sigma^*$ such that, say, $(x, t + \sigma^* - \sigma'') \succ (y, s + \sigma^* - \sigma'')$

with $t+\sigma^*-\sigma'' \geq 0$. By continuity and impatience there exist $t' > t$ such that $(x, t'+\sigma^*-\sigma'') \sim (y, s+\sigma^*-\sigma'')$. Then, since $\succcurlyeq$ exhibits constant impatience it follows that $(x, t'+\sigma^*-\sigma''+\gamma) \sim (y, s+\sigma^*-\sigma''+\gamma)$ for all $\gamma > 0$. Let $\gamma = \sigma'' > 0$. Then we obtain $(x, t'+\sigma^*) \sim (y, s+\sigma^*)$. Since $t' > t$ it implies by impatience that $(x, t'+\sigma^*) \prec (x, t+\sigma^*)$. Hence, $(x, t+\sigma^*) \succ (y, s+\sigma^*)$, a contradiction.

Suppose that $t < s$ and $\succcurlyeq$ exhibits strictly DI. It follows that $(x, t+\sigma^*+\alpha) \prec (y, s+\sigma^*+\alpha)$ for all $\alpha > 0$, or $(x, t+\sigma) \prec (y, s+\sigma)$ for all $\sigma > \sigma^*$. We now need to show that $(x, t+\sigma) \succ (y, s+\sigma)$ for all $\sigma < \sigma^*$, or $(x, t+\sigma^*-\sigma') \succ (y, s+\sigma^*-\sigma')$ for all $\sigma' > \sigma^*$ such that $t+\sigma^*-\sigma' \geq 0$. The proof is by contradiction. Suppose there exist $\sigma'' > \sigma^*$ such that $(x, t+\sigma^*-\sigma'') \preccurlyeq (y, s+\sigma^*-\sigma'')$ and $t+\sigma^*-\sigma'' \geq 0$.

First consider $(x, t+\sigma^*-\sigma'') \sim (y, s+\sigma^*-\sigma'')$ with $\sigma'' > \sigma^*$ and $t+\sigma^*-\sigma'' \geq 0$. Then, since $\succcurlyeq$ satisfy strictly DI it follows that $(x, t+\sigma^*-\sigma''+\gamma) \prec (y, s+\sigma^*-\sigma''+\gamma)$ for all $\gamma > 0$. Let $\gamma = \sigma'' > 0$. Then we have $(x, t+\sigma^*) \prec (y, s+\sigma^*)$, which is a contradiction. Secondly, consider $(x, t+\sigma^*-\sigma'') \prec (y, s+\sigma^*-\sigma'')$ with $\sigma'' > \sigma^*$ and $t+\sigma^*-\sigma'' \geq 0$. It follows by continuity and impatience that there exist $s' > s$ such that $(x, t+\sigma^*-\sigma'') \sim (y, s'+\sigma^*-\sigma'')$. Hence, since $\succcurlyeq$ exhibit strictly DI it implies that $(x, t+\sigma^*-\sigma''+\gamma) \prec (y, s'+\sigma^*-\sigma''+\gamma)$ for all $\gamma > 0$ with $\sigma'' > \sigma^*$ and $t+\sigma^*-\sigma'' \geq 0$. Let $\gamma = \sigma''$. Then we have $(x, t+\sigma^*) \prec (y, s'+\sigma^*)$. Since $s' > s$ it follows by impatience that $(y, s'+\sigma^*) \prec (y, s+\sigma^*)$, therefore, $(x, t+\sigma^*) \prec (y, s+\sigma^*)$, a contradiction.

If we assume that $\succcurlyeq$ exhibits strictly II, the proof is analogous. $\square$

While the assumption of a DU representation is not necessary for the *"If"* part of the proof, it is essential for our proof of the *"Only If"* part. The necessity of a DU representation for the *"Only If"* result of Lemma 4.11 remains an open question.

It was demonstrated in the working paper [7], that a strictly increasing time preference rate corresponds to strictly II, while a strictly decreasing time preference rate corresponds to strictly DI. The proof is along the lines of Lemma 3.15 in Chapter 3. We use this result to prove the following proposition.

Proposition 4.12. *Suppose that $\succcurlyeq$ has a DU representation (u, D).*

- *If $D(t) = (1+ct)e^{-rt}$, where $r \geq c \geq 0$ and $r > 0$, then $\succcurlyeq$ exhibits strictly II when $c > 0$ and $\succcurlyeq$ exhibits stationarity when $c = 0$.*

- *If $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, where $a, b, c > 0$, $a \leq b/c+1$, then $\succcurlyeq$ exhibits strictly DI when $a < 1$, strictly II when $1 < a \leq b/c+1$ and stationarity when $a = 1$.*

Proof. (a) Linear times exponential. Assume first that $c > 0$. Since $D'(t) = e^{-rt}(c - r - crt)$, the time preference rate[4] is:

$$-\frac{D'(t)}{D(t)} = \frac{e^{-rt}(crt - c + r)}{(1+ct)e^{-rt}} = \frac{r(1+ct) - c}{1+ct} = r - \frac{c}{1+ct}.$$

The derivative of time preference rate is $c^2(1+ct)^{-2} > 0$, therefore, linear times exponential discount function exhibits strictly increasing impatience. Otherwise, if $c = 0$, then $D(t) = e^{-rt}$ and the preferences $\succcurlyeq$ exhibit stationarity (see, for example, [32]).

(b) Sums of exponentials. The time preference rate is

$$\begin{aligned}-\frac{D'(t)}{D(t)} &= \frac{abe^{-bt} + (1-a)(b+c)e^{-(b+c)t}}{ae^{-bt} + (1-a)e^{-(b+c)t}} = \frac{e^{-bt}\left[ab + (1-a)(b+c)e^{-ct}\right]}{e^{-bt}\left[a + (1-a)e^{-ct}\right]} \\ &= \frac{ab + (1-a)(b+c)e^{-ct}}{a + (1-a)e^{-ct}}.\end{aligned}$$

The derivative of time preference rate is

$$\begin{aligned}&\left[\frac{ab + (1-a)(b+c)e^{-ct}}{a + (1-a)e^{-ct}}\right]' \\ =\ &\frac{-c(1-a)(b+c)e^{-ct}\left[a + (1-a)e^{-ct}\right] + \left[ab + (1-a)(b+c)e^{-ct}\right]c(1-a)e^{-ct}}{\left[a + (1-a)e^{-ct}\right]^2}.\end{aligned}$$

The sign of the derivative depends on the sign of the numerator of this expression:

$$Q(t) = -c(1-a)(b+c)e^{-ct}\left[a + (1-a)e^{-ct}\right] + \left[ab + (1-a)(b+c)e^{-ct}\right]c(1-a)e^{-ct}.$$

Simplifying $Q(t)$:

$$Q(t) = c(1-a)e^{-ct}\left[ab + (1-a)(b+c)e^{-ct} - (b+c)(a + (1-a)e^{-ct})\right] = c^2e^{-ct}a(a-1).$$

Recall that $a > 0$ and $a \le b/c + 1$. Therefore, $Q(t) = 0$ if $a = 1$, $Q(t)$ is strictly negative if $0 < a < 1$ and $Q(t)$ is strictly positive if $1 < a \le b/c + 1$ and $a \ne 1$. Hence, the time preference rate is constant if $a = 1$, strictly decreasingif $0 < a < 1$ and strictly increasing if $1 < a \le b/c + 1$. This in turn implies that $\succcurlyeq$ exhibit stationarity if $a = 1$, strictly DI if $0 < a < 1$ and strictly II if $1 < a \le b/c + 1$. □

A linear times exponential discount function and two sum of exponentials discount functions, with their associated rates of time preference, are illustrated in Figure 4.2.

[4] See also Lemma 3.15 in Section 3.2.2 of Chapter 3.

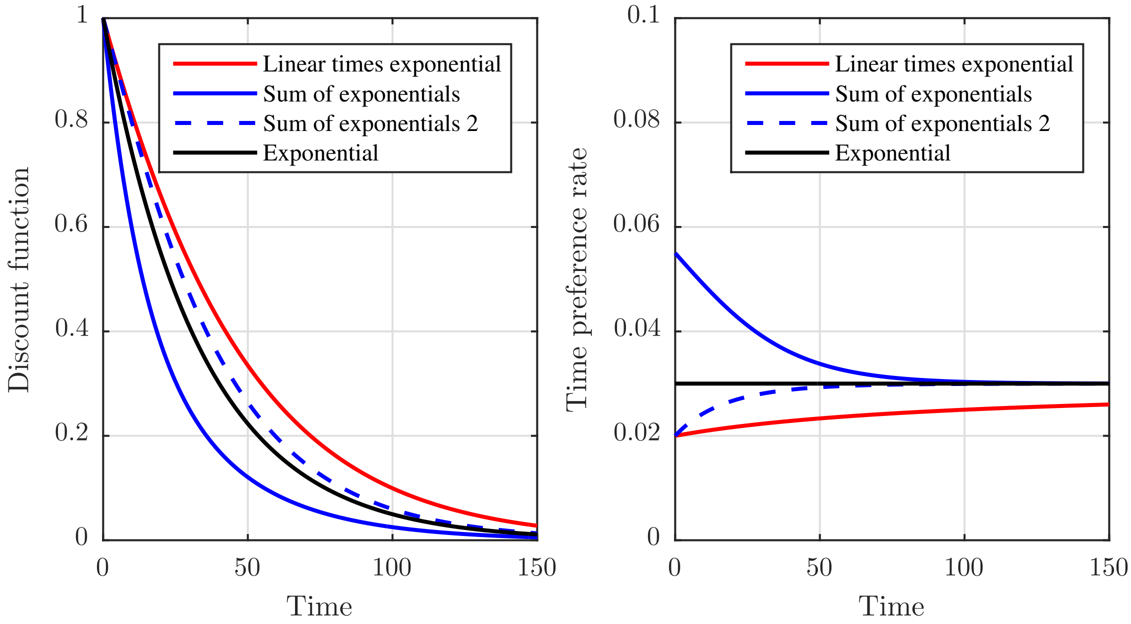

Figure 4.2: Linear times exponential discount function $D(t) = (1 + 0.01t)e^{-0.03t}$, exponential discount function $D(t) = e^{-0.03t}$, sum of exponentials $D(t) = 0.5e^{0.03t} + 0.5e^{0.08t}$, sum of exponentials $D(t) = 1.2e^{0.03t} - 0.2e^{0.08t}$ and their associated rates of time preference

It is worth mentioning that Bell's [12] definitions of the terms "decreasing impatience" and "increasing impatience" are different from the ones used here. Bell's definitions are given below:

Definition 4.13 ([12])**.** *Let $\succcurlyeq$ on $\mathcal{A}_1$ have a DU representation with a discount function D. Then we say that preferences $\succcurlyeq$ exhibit DI* [II*] if*

$$D(s+t) > [<]\ D(s)D(t) \text{ for any } s, t > 0.$$

Note that Bell's [12] DI* (II*) corresponds to strict log-superadditivity (log-subadditivity) of D. Obviously, strict log-superadditivity (strict log-subadditivity) is a special case of strict log-convexity (strict log-concavity). Therefore, Bell's definitions of DI* and II* are weaker properties than our strictly DI and strictly II. Bell [12, Proposition 8] specifies the parameter values for a sum of exponentials discount function such that $\succcurlyeq$ exhibit DI*/II*:

Proposition 4.14 ([12, Proposition 8])**.** *Let $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, where $a, b, c > 0$ and $a \leq 1 + b/c$. Then it is DI* if $a < 1$ and II* if $1 < a \leq 1 + b/c$.*[5]

[5]Bell [12, Proposition 8] uses a strict inequality $a < 1 + b/c$ to guarantee that $D'(t) < 0$ when $t = 0$. However, if $D'(t) = 0$ when $t = 0$ and $D'(t) < 0$ for all $t > 0$, then $D(t)$ is strictly decreasing on $[0, \infty)$. Therefore, we allow $a \leq 1 + b/c$, since this weak inequality is consistent with the properties of a discount function.

Comparing Proposition 4.14 to Proposition 4.11, it is easy to see that strictly II and II* have the same implications for the parameters for a sum of exponentials discount function (and similarly for strictly DI and DI*). It is also straightforward to observe that the restrictions imposed by the properties of strictly II and II* on the parameters for a linear times exponential discount function coincide.

4.2.3 Representation of preferences with the one-switch property

We first show that constant impatience is equivalent to the zero-switch property for preferences with a DU representation:

Definition 4.15. *We say that $\succcurlyeq$ on $\mathcal{A}$ exhibit the zero-switch property if for every pair $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ the ranking of $(\mathbf{x}, \mathbf{t} + \sigma)$ and $(\mathbf{y}, \mathbf{s} + \sigma)$ is independent of σ.*

It follows from this definition that if $\succcurlyeq$ on $\mathcal{A}$ exhibit the zero-switch property, then for any $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$, if there exists $\hat{\sigma} \geq 0$ such that $(\mathbf{x}, \mathbf{t} + \hat{\sigma}) \sim (\mathbf{y}, \mathbf{s} + \hat{\sigma})$, then $(\mathbf{x}, \mathbf{t} + \sigma) \sim (\mathbf{y}, \mathbf{s} + \sigma)$ for any $\sigma \geq 0$.

The following proposition amends [12, Proposition 8].

Proposition 4.16. *Let $\succcurlyeq$ on $\mathcal{A}$ have a DU representation (u, D).Then $\succcurlyeq$ exhibit the one-switch property only if $D(t)$ has one of the following forms:*

- $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$*, with* $a, b, c > 0$ *and* $a \leq b/c + 1$*,*
- $D(t) = (1 + ct)e^{-rt}$*, where* $r \geq c \geq 0$ *and* $r > 0$*.*

We adapt Bell's [12, 13] method of proof for the risk (expected utility) context to our time preference framework. The required adaptation is non-trivial. The main reason is that probabilities sum up to one, whereas utilities of outcomes do not. In the original Bell proof [12] this property of probabilities was used to obtain a system of two equations in two variables. The conditions under which the solutions of this system exist are well-known. In the time preference setting, we use Lemma 4.11 to obtain two sequences of dated outcomes that are indifferent at two different points of delay. As a result of which we obtain a homogeneous second order linear difference equation. The solutions of this equation extended to continuous time give us three types of functions. We further show that only linear times exponential and sums of exponentials (with suitable parameter restrictions) satisfy the one-switch property and the properties of a discount function (using Proposition 4.4). It is also worth mentioning that in the risk setting Bell [12, 13] obtains a third order difference equation, rather than the second order difference equation that we obtain in the time preference framework.

Proof. Fix some $\sigma > 0$. Suppose we can find $u(\alpha), u(\beta), u(\gamma) > 0$ such that

$$u(\alpha) + u(\beta)D(2\sigma) = u(\gamma)D(\sigma) \text{ and } u(\alpha)D(\sigma) + u(\beta)D(3\sigma) = u(\gamma)D(2\sigma) \tag{4.24}$$

Then, since $\succcurlyeq$ has a DU representation, it implies that

$$(\mathbf{x}, \mathbf{t}) \sim (\mathbf{y}, \mathbf{s}) \text{ and } (\mathbf{x}, \mathbf{t} + \sigma) \sim (\mathbf{y}, \mathbf{s} + \sigma),$$

where $(\mathbf{x}, \mathbf{t}) = ((\alpha, \beta), (0, 2\sigma))$, $(\mathbf{y}, \mathbf{s}) = (\gamma, \sigma)$ and $(\mathbf{x}, \mathbf{t} + \sigma) = ((\alpha, \beta), (\sigma, 3\sigma))$, $(\mathbf{y}, \mathbf{s} + \sigma) = (\gamma, 2\sigma)$.

We first show that for any $D(\sigma), D(2\sigma), D(3\sigma)$ we can always find $u(\alpha), u(\beta), u(\gamma) > 0$ such that (4.24) holds. The first equation of 4.24 implies $u(\alpha) = u(\gamma)D(\sigma) - u(\beta)D(2\sigma)$, which we may substitute into the second equation as follows:

$$\left(u(\gamma)D(\sigma) - u(\beta)D(2\sigma)\right) D(\sigma) + u(\beta)D(3\sigma) = u(\gamma)D(2\sigma).$$

Simplifying

$$u(\gamma)(D(\sigma)^2 - D(2\sigma)) + u(\beta) \left(D(3\sigma) - D(\sigma)D(2\sigma)\right) = 0. \tag{4.25}$$

Since preferences $\succcurlyeq$ satisfy the one-switch property they will also satisfy the one-switch property for dated outcomes, and hence it follows by Lemma 4.11 that $\succcurlyeq$ exhibits strictly DI, strictly II or stationarity. Assume first that $\succcurlyeq$ exhibits strictly DI. Then, by Proposition 4.8 we have

$$\frac{1}{D(\sigma)} > \frac{D(\sigma)}{D(2\sigma)} > \frac{D(2\sigma)}{D(3\sigma)} > \cdots$$

Therefore, $D(\sigma)^2 - D(2\sigma) < 0$ and $D(3\sigma) - D(\sigma)D(2\sigma) > 0$, and hence, by continuity it is always possible to find $u(\beta), u(\gamma) > 0$ such that (4.25) holds.

Since $\frac{u(\beta)}{u(\gamma)} = \frac{D(2\sigma) - D(\sigma)^2}{D(3\sigma) - D(\sigma)D(2\sigma)}$, we have

$$\frac{u(\beta)}{u(\gamma)} - \frac{D(\sigma)}{D(2\sigma)} = \frac{D(2\sigma) - D(\sigma)^2}{D(3\sigma) - D(\sigma)D(2\sigma)} - \frac{D(\sigma)}{D(2\sigma)} = \frac{D(2\sigma)^2 - D(\sigma)D(3\sigma)}{D(2\sigma)\left(D(3\sigma) - D(\sigma)D(2\sigma)\right)} < 0.$$

Hence, $u(\beta)D(2\sigma) - u(\gamma)D(\sigma) < 0$, which implies that

$$u(\alpha) = u(\gamma)D(\sigma) - u(\beta)D(2\sigma) > 0.$$

If $\succcurlyeq$ exhibits strictly II the argument is analogous with the inequalities reversed.

In the case when $\succcurlyeq$ exhibits stationarity, we have $D(\sigma)^2 - D(2\sigma) = 0$ and $D(3\sigma) -$

$D(\sigma)D(2\sigma) = 0$, therefore (4.25) holds for any $u(\beta), u(\gamma)$. Hence, we can choose $u(\beta), u(\gamma) > 0$ such that $u(\alpha) = u(\gamma)D(\sigma) - u(\beta)D(2\sigma) > 0$.

Therefore, we have found two sequences of dated outcomes $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ such that

$$(\mathbf{x}, \mathbf{t} + t) \sim (\mathbf{y}, \mathbf{s} + t), \text{ where } t = 0, \sigma.$$

Then, since $\succcurlyeq$ satisfies the one-switch property it must be true that

$$(\mathbf{x}, \mathbf{t} + t) \sim (\mathbf{y}, \mathbf{s} + t) \text{ for any } t \geq 0.$$

In particular, $u(\alpha)D(t) + u(\beta)D(t + 2\sigma) = u(\gamma)D(t + \sigma)$.

Since $u(\gamma) \neq 0$, we can write $D(t + 2\sigma) + aD(t + \sigma) + bD(t) = 0$, where $a = -u(\gamma)/u(\beta)$, $b = u(\alpha)/u(\beta)$.

For some $\sigma > 0$ and some $t_0 \geq 0$ let $D_n^{(t_0,\sigma)} = D(t_0 + n\sigma)$, $n \in \mathbb{N}_0$. We then obtain a homogeneous second order linear difference equation

$$D_{n+2}^{(t_0,\sigma)} + aD_{n+1}^{(t_0,\sigma)} + bD_n^{(t_0,\sigma)} = 0. \tag{4.26}$$

It is well known (see, for example, [55]) that this equation has three types of solutions, which are derived from the characteristic equation: $z^2 + az + b = 0$. These three solutions are as follows:

Solution 1. If $a^2 - 4b > 0$, then there are two distinct real roots denoted as z_1 and z_2. In this case the two linearly independent solutions to (4.26) are z_1^n and z_2^n, where $n \in \mathbb{N}_0$. The general solution is $D_n^{(t_0,\sigma)} = c_1 z_1^n + c_2 z_2^n$, where $c_1, c_2 = const.$

Solution 2. If $a^2 - 4b = 0$, then the roots coincide so $z_1 = z_2 = z_0$. In this case the two linearly independent solutions to (4.26) are z_0^n and nz_0^n. The general solution is $D_n^{(t_0,\sigma)} = (c_1 + c_2 n)\, z_0^n$, where $c_1, c_2 = const.$

Solution 3. If $a^2 - 4b < 0$, the roots are complex

$$z_\pm \;=\; x \pm iy \;=\; r\,(\cos\theta \pm i \sin\theta) \;=\; re^{\pm i\theta},$$

where $y > 0$, $r = \sqrt{x^2 + y^2}$, $\cos\theta = x/r$ and $\sin\theta = y/r$ with $\theta \in (0, \pi)$ (since $y > 0$). Then the two linearly independent solutions to (4.26) are $Re z_+^n = r^n \cos n\theta$ and $Im z_+^n = r^n \sin n\theta$. The general solution is

$$D_n^{(t_0,\sigma)} = r^n \left[c_1 \cos n\theta + c_2 \sin n\theta\right],$$

where $c_1, c_2 = const.$

Note that by letting $C = \sqrt{c_1^2 + c_2^2}$, $\cos\omega = \frac{c_1}{C}$, $\sin\omega = \frac{c_2}{C}$ and $\omega = \tan^{-1}(\frac{c_2}{c_1})$, Solution 3 can be rewritten as follows:

$$D_n^{(t_0,\sigma)} = r^n\left[c_1 \cos n\theta + c_2 \sin n\theta\right] = Cr^n\left[\cos\omega\cos n\theta + \sin\omega \sin n\theta\right] = Cr^n\cos(n\theta - \omega).$$

Recall that $\theta \in (0, \pi)$. Therefore, Solution 3 can be excluded because it implies multiple changes of sign (it does not satisfy monotonicity).

Note that equation (4.26) holds for arbitrary $\sigma > 0$ and arbitrary $t_0 \geq 0$, though the roots z_1 and z_2 and the constants c_1 and c_2 may depend, in a continuous fashion, on t_0 and σ. Bell [13, 12] argues that D must therefore satisfy the corresponding limit of one of these solutions, as $\sigma \to 0$.[6] The solutions of (4.26) converge, respectively to

Solution 1. (Sum of exponentials): $D(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}$, where $r_1 \neq r_2$,

Solution 2. (Linear times exponential): $D(t) = (c_1 + c_2 t)e^{r_0 t}$.

By Proposition 4.4 it follows that:

Solution 1. $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, where $a, b, c > 0$, and $a \leq b/c + 1$;

Solution 2. $D(t) = (1 + ct)e^{-rt}$, where $r \geq c \geq 0$ and $r > 0$. Note that this solution includes the exponential discount function as a special case. That is, if $c = 0$, then $D(t) = e^{-rt}$, where $r > 0$.

□

The following corollary summarizes Proposition 4.5 and Proposition 4.16.

Corollary 4.17. *Let $\succcurlyeq$ on $\mathcal{A}$ have a DU representation (u, D). Then $\succcurlyeq$ exhibits the one-switch property if and only if $D(t)$ has one of the following forms:*

- $D(t) = ae^{-bt} + (1-a)e^{-(b+c)t}$, *with $a, b, c > 0$ and $a \leq b/c + 1$,*
- $D(t) = (1 + ct)e^{-rt}$, *where $r \geq c \geq 0$ and $r > 0$.*

[6] Romanian mathematician Radó [67] proved a more general result which was recently extended to multidimensional case in [5]. Radó [67, Theorem 2] proves that the set of continuous functions $f : \mathbb{R} \to \mathbb{R}$, which satisfy the equation $a_0(\sigma)f(t) + a_1(\sigma)f(t+\sigma) + \ldots + a_n(\sigma)f(t+n\sigma) = 0$ with continuous functions $a_i(\sigma)\colon (0, W) \to \mathbb{R}$, where $W > 0$ and $i = 0, \ldots, n$ with $a_n(\sigma) \neq 0$, coincides with the set of functions $f\colon \mathbb{R} \to \mathbb{R}$ which satisfy a linear differential equation $A_0 f + A_1 f' + \ldots + A_n f^{(n)} = 0$ for some real coefficients $A_0, \ldots, A_n$. When $n = 2$ it is well-known that the solutions to such a differential equation coincide with the limits of our three solution types.

4.3 The weak one-switch property

4.3.1 The weak one-switch property and mixtures of sequences of dated outcomes

The one-switch property is equivalent to the weak one-switch property in the risk setting where preferences over lotteries have an expected utility representation [12]. This equivalence follows from [12, Lemma 3], where mixture linearity and other properties of expected utility are used to show that if two lotteries are indifferent at two wealth levels then they should be indifferent for any wealth level. In the intertemporal framework of this chapter a direct adaptation of the proof of this equivalence is not possible, even if we assume that preferences have a DU representation, because a DU representation is, in general, not mixture linear. However, in Chapter 2 we introduced an Anscombe and Aumann (AA) setting [8] with preferences over streams of consumption lotteries. It turns out that another benefit of working in an environment like AA [8] is that it is possible to establish the equivalence of the weak one-switch property and the one-switch property for time preferences. We have to adapt the AA framework to continuous time for this purpose.

Assume that X is a **mixture set** (as defined in Section 2.1 of Chapter 2); that is, for every $x, y \in X$ and every $\lambda, \mu \in [0,1]$, there exists $x\lambda y \in X$ satisfying:

- $x1y = x$,
- $x\lambda y = y(1-\lambda)x$,
- $(x\mu y)\lambda y = x(\lambda\mu)y$.

Recall that in Chapter 2 a "neutral" (status quo) outcome was denoted as x_0. We change the notation here and assume that X contains a "neutral" outcome, denoted by 0. We can think of X as a set of lotteries with monetary outcomes, and 0 corresponds to the lottery which pays 0 with certainty.

We next introduce a mixture operation for sequences of dated outcomes, analogous to the AA mixing operation. To do so, we recall that the neutral outcome obtains at any date not specified in the sequence. First, define

$$(\mathbf{x}, \mathbf{t})\lambda(\mathbf{y}, \mathbf{t}) = (\mathbf{x}\lambda\mathbf{y}, \mathbf{t}) \text{ for any } (\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{t}) \in \mathcal{A} \text{ and all } \lambda \in [0,1], \tag{4.27}$$

where $\mathbf{x}\lambda\mathbf{y}$ is defined the usual way (see Section 2.1 of Chapter 2).

Let $(\mathbf{x}, \mathbf{t}) \in \mathcal{A}$ with $\mathbf{t} = (t_1, t_2, \ldots, t_n)$ and let $\mathbf{s} = (s_1, s_2, \ldots, s_m)$. Define $\mathbf{t}|\mathbf{s} = \mathbf{l} = (l_1, l_2, \ldots, l_k)$, where $l_1 < l_2 < \ldots < l_k$ and $\{l_1, l_2, \ldots, l_k\} = \{t_1, t_2, \ldots, t_n\} \cup$

$\{s_1, s_2, \ldots, s_m\}$. An example of concatenation procedure for time vectors $\mathbf{t} = (t_1, t_2, t_3)$ and $\mathbf{s} = (s_1, s_2, s_3, s_4)$ is given in Figure 4.3.

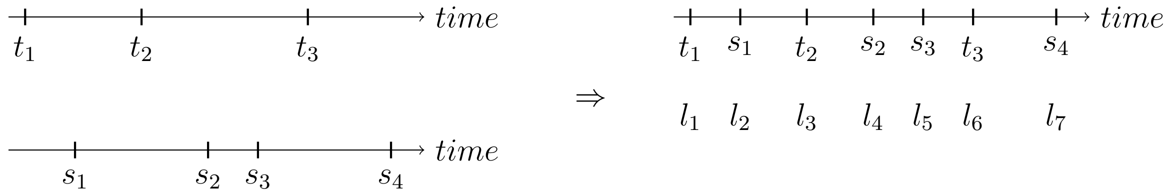

Figure 4.3: Concatenation of $\mathbf{t} = (t_1, t_2, t_3)$ and $\mathbf{s} = (s_1, s_2, s_3, s_4)$

For any $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ and any $\lambda \in [0, 1]$ define the mixture operation as follows:

$$(\mathbf{x}, \mathbf{t})\lambda(\mathbf{y}, \mathbf{s}) = (\mathbf{z}, \mathbf{t}|\mathbf{s})\lambda(\mathbf{z}', \mathbf{s}|\mathbf{t}), \tag{4.28}$$

where $\mathbf{z}$ is defined so that if $l_j = t_i$, then $z_j = x_i$, otherwise $z_j = 0$, and $\mathbf{z}'$ is defined so that if $l_j = s_i$, then $z'_j = y_i$, otherwise $z'_j = 0$. Note that $(\mathbf{x}, \mathbf{t})$ and $(\mathbf{z}, \mathbf{t}|\mathbf{s})$ are identical sequences of dated outcomes, and likewise $(\mathbf{y}, \mathbf{s})$ and $(\mathbf{z}', \mathbf{s}|\mathbf{t})$ are identical sequences. Figure 4.4 illustrates the transformation of the sequence $(\mathbf{x}, \mathbf{t}) = ((x_1, x_2, x_3), (t_1, t_2, t_3))$ to the sequence $(\mathbf{z}, \mathbf{t}|\mathbf{s})$ and the sequence $(\mathbf{y}, \mathbf{s}) = ((y_1, y_2, y_3, y_4), (s_1, s_2, s_3, s_4))$ to the sequence $(\mathbf{z}', \mathbf{s}|\mathbf{t})$. Note that $\mathbf{t}|\mathbf{s} = \mathbf{s}|\mathbf{t}$.

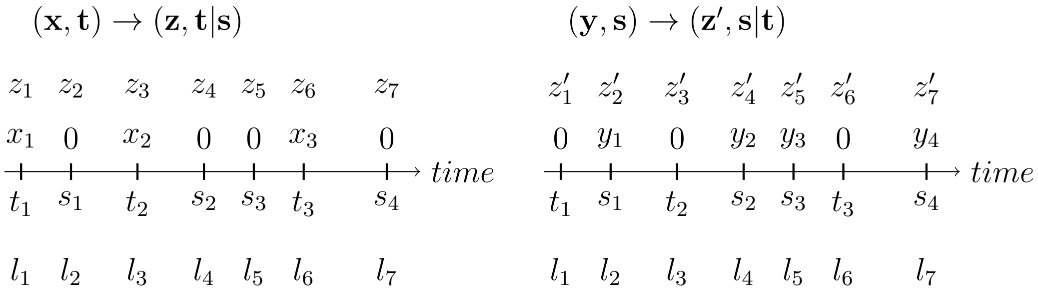

Figure 4.4: Transformation of $(\mathbf{x}, \mathbf{t}) = ((x_1, x_2, x_3), (t_1, t_2, t_3))$ to $(\mathbf{z}, \mathbf{t}|\mathbf{s})$ and $(\mathbf{y}, \mathbf{s}) = ((y_1, y_2, y_3, y_4), (s_1, s_2, s_3, s_4))$ to $(\mathbf{z}', \mathbf{s}|\mathbf{t})$

It is not hard to see that $\mathcal{A}$ is a mixture set.

The utility function $u\colon X \to \mathbb{R}$ is called *mixture linear* if for every $x, y \in X$ we have $u(x\lambda y) = \lambda u(x) + (1 - \lambda)u(y)$.

We say that preferences in this environment have *a DEU (discounted expected utility) representation* if they have a DU representation (u, D) in which u is mixture linear on X. We next show that the induced utility U on sequences of dated outcomes is mixture linear on $\mathcal{A}$. It follows from (4.28) that

$$U((\mathbf{x}, \mathbf{t})\lambda(\mathbf{y}, \mathbf{s})) \;=\; U((\mathbf{z}, \mathbf{t}|\mathbf{s})\lambda(\mathbf{z}', \mathbf{s}|\mathbf{t})) \;=\; U((\mathbf{z}, \mathbf{l})\lambda(\mathbf{z}', \mathbf{l})) \;=\; U(\mathbf{z}\lambda\mathbf{z}', \mathbf{l}), \tag{4.29}$$

where $\mathbf{l} = \mathbf{t}|\mathbf{s} = \mathbf{s}|\mathbf{t}$. Since u is mixture linear it follows that

$$\begin{aligned} U(\mathbf{z}\lambda\mathbf{z}', \mathbf{l}) &= \lambda U(\mathbf{z}, \mathbf{l}) + (1-\lambda)U(\mathbf{z}', \mathbf{l}) = \lambda U(\mathbf{z}, \mathbf{t}|\mathbf{s}) + (1-\lambda)U(\mathbf{z}', \mathbf{s}|\mathbf{t}) \\ &= \lambda U(\mathbf{x}, \mathbf{t}) + (1-\lambda)U(\mathbf{y}, \mathbf{s}). \end{aligned}$$

Hence, U is mixture linear on $\mathcal{A}$.

It is worth mentioning that the existence of a suitable DEU axiomatization remains an open question. It is natural to assume that Fishburn and Rubinstein' s axioms [32] should be satisfied for $\succcurlyeq$ restricted to $\mathcal{A}_1$ with X being a mixture set.[7] However, it is beyond the scope of the present chapter to address this issue.

The following lemma is a part of [12, Lemma 3]. The proof is given for completeness.

Lemma 4.18 (cf., [12])**.** *Let preference $\succcurlyeq$ on $\mathcal{A}$ have a DU representation. If $\succcurlyeq$ on $\mathcal{A}$ exhibit the weak one-switch property, then for any $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ if there exist σ_1, σ_2 such that $\sigma_1 < \sigma_2$ and $(\mathbf{x}, \mathbf{t} + \sigma_1) \sim (\mathbf{y}, \mathbf{s} + \sigma_1)$, and $(\mathbf{x}, \mathbf{t} + \sigma_2) \sim (\mathbf{y}, \mathbf{s} + \sigma_2)$, then $(\mathbf{x}, \mathbf{t} + \sigma) \sim (\mathbf{y}, \mathbf{s} + \sigma)$ for any $\sigma \in (\sigma_1, \sigma_2)$.*

Proof. Suppose $(\mathbf{x}, \mathbf{t}), (\mathbf{y}, \mathbf{s}) \in \mathcal{A}$ and σ_1, σ_2 are such that $\sigma_1 < \sigma_2$ with

$$\begin{aligned} (\mathbf{x}, \mathbf{t} + \sigma_1) &\sim (\mathbf{y}, \mathbf{s} + \sigma_1), \\ (\mathbf{x}, \mathbf{t} + \sigma_2) &\sim (\mathbf{y}, \mathbf{s} + \sigma_2). \end{aligned}$$

We need to show that $(\mathbf{x}, \mathbf{t} + \sigma) \sim (\mathbf{y}, \mathbf{s} + \sigma)$ for any $\sigma \in (\sigma_1, \sigma_2)$. The proof is by contradiction. Assume that we can have $\hat{\sigma} \in (\sigma_1, \sigma_2)$ such that $(\mathbf{x}, \mathbf{t} + \hat{\sigma}) \succ (\mathbf{y}, \mathbf{s} + \hat{\sigma})$. Consider $\mathbf{y}' = \mathbf{y} + \varepsilon$, where $\varepsilon > 0$ is sufficiently small so that $(\mathbf{x}, \mathbf{t} + \hat{\sigma}) \succ (\mathbf{y}', \mathbf{s} + \hat{\sigma})$ (by continuity, such ε can always be found). However, $(\mathbf{x}, \mathbf{t} + \sigma_1) \prec (\mathbf{y}', \mathbf{s} + \sigma_1)$ and $(\mathbf{x}, \mathbf{t} + \sigma_2) \prec (\mathbf{y}', \mathbf{s} + \sigma_2)$, which implies a double switch. We obtained the desired contradiction. The case $(\mathbf{x}, \mathbf{t} + \hat{\sigma}) \prec (\mathbf{y}, \mathbf{s} + \hat{\sigma})$ is similar. Therefore, $(\mathbf{x}, \mathbf{t} + \sigma) \sim (\mathbf{y}, \mathbf{s} + \sigma)$ for any $\sigma \in (\sigma_1, \sigma_2)$. □

Lemma 4.18 implies the following corollary:

Corollary 4.19. *The weak one-switch property implies that $\{\sigma \mid \Delta(\sigma) = 0\}$ is a closed (possibly empty) interval.*

In the next proposition the mixtures of sequences of dated outcomes will be used. It adapts [12, Lemma 3] to the time preference setting. Proposition 4.20 is proved by contradiction. We first need to find two sequences of dated outcomes such that

[7] For example, Bleichrodt et al. [18] assume the existence of DU representation for $\succcurlyeq$ on $\mathcal{A}_1$, where X is not necessarily restricted to reals. Similarly, Rohde [70] applies the DU representation to $\succcurlyeq$ on $\mathcal{A}_1$ requiring only that X is a connected topological space which contains a "neutral" outcome.

their DEU difference changes its sign as the delay increases. Using Corollary 4.19 we then consider two cases depending on whether the just obtained DEU difference equals zero at a unique point or on the interval. A double switch (contradiction to the weak one-switch property) can be obtained in each case by introducing suitable mixtures of sequences of dated outcomes. The illustrations to support the proof are given in the Appendix.

Proposition 4.20 (cf., [12, Lemma 3]). *Let preference $\succcurlyeq$ on $\mathcal{A}$ have a DEU representation. If $\succcurlyeq$ exhibits the weak one-switch property, then for any $(\mathbf{x},\mathbf{t}),(\mathbf{y},\mathbf{s}) \in \mathcal{A}$ if $(\mathbf{x},\mathbf{t}+\sigma_1) \sim (\mathbf{y},\mathbf{s}+\sigma_1)$ and $(\mathbf{x},\mathbf{t}+\sigma_2) \sim (\mathbf{y},\mathbf{s}+\sigma_2)$ for some $\sigma_1 \geq 0$ and some $\sigma_2 > \sigma_1$, then $(\mathbf{x},\mathbf{t}+\sigma) \sim (\mathbf{y},\mathbf{s}+\sigma)$ for any $\sigma \geq 0$.*

Proof. Suppose $0 \leq \sigma_1 < \sigma_2$ and

$$(\mathbf{x},\mathbf{t}+\sigma_1) \sim (\mathbf{y},\mathbf{s}+\sigma_1),$$
$$(\mathbf{x},\mathbf{t}+\sigma_2) \sim (\mathbf{y},\mathbf{s}+\sigma_2).$$

By Lemma 4.18 it follows that

$$(\mathbf{x},\mathbf{t}+\sigma) \sim (\mathbf{y},\mathbf{s}+\sigma) \text{ for any } \sigma \in (\sigma_1,\sigma_2). \tag{4.30}$$

The weak one-switch property implies weak preference in one direction above σ_2 and in the other direction below σ_1. We assume that $(\mathbf{x},\mathbf{t}+\sigma)$ is weakly preferred to $(\mathbf{y},\mathbf{s}+\sigma)$ above σ_2. The other case can be treated similarly.

Suppose that (4.30) is satisfied but

$$(\mathbf{x},\mathbf{t}+\sigma') \succ (\mathbf{y},\mathbf{s}+\sigma') \text{ for some } \sigma' > \sigma_2. \tag{4.31}$$

We will prove that a contradiction follows.

Let $\Delta_1(\sigma) = U(\mathbf{x},\mathbf{t}+\sigma) - U(\mathbf{y},\mathbf{s}+\sigma)$. Therefore,

$$\Delta_1(\sigma) = 0, \text{ if } \sigma \in [\sigma_1,\sigma_2], \text{ and } \Delta_1(\sigma') > 0.$$

Given the DEU representation we can always find $a, b \in X$ and $q > p$ such that

$$\Delta_2(\sigma) = U(a, q+\sigma) - U(b, p+\sigma) < 0 \text{ for all } \sigma.$$

Let $(\mathbf{z_1},\mathbf{t_1}) = (a,q)\lambda(\mathbf{x},\mathbf{t})$, where $\lambda \in (0,1)$. Analogously, define $(\mathbf{z_2},\mathbf{t_2}) = (b,p)\lambda(\mathbf{y},\mathbf{s})$. Consider the function

$$\Delta_3(\sigma) \;=\; U(\mathbf{z_1},\mathbf{t_1}+\sigma) - U(\mathbf{z_2},\mathbf{t_2}+\sigma)$$

$$
\begin{aligned}
&= \lambda\left(U(a,q+\sigma)-U(b,p+\sigma)\right)+(1-\lambda)\left(U(\mathbf{x},\mathbf{t}+\sigma)-U(\mathbf{y},\mathbf{s}+\sigma)\right)\\
&= \lambda\Delta_2(\sigma)+(1-\lambda)\Delta_1(\sigma).
\end{aligned}
$$

Then for λ sufficiently small we have (see Figure A.1 in the Appendix[8]):

$$\Delta_3(\sigma_1)<0, \Delta_3(\sigma_2)<0, \text{ and } \Delta_3(\sigma')>0.$$

Therefore, by continuity, there must exist at least one $\sigma^* \in (\sigma_2,\sigma')$ such that $\Delta_3(\sigma^*)=0$. By Corollary 4.19 there are two possible cases: either σ^* is unique, or $\Delta_3(\sigma)=0$ for $\sigma\in[\sigma_1^*,\sigma_2^*]\subset(\sigma_2,\sigma')$ (see Figure A.2 in the Appendix).

Case 1. Assume first that σ^* is unique; i.e., (using the one-switch property) $\Delta_3(\sigma)<0$ if $\sigma<\sigma^*$ and $\Delta_3(\sigma)>0$ if $\sigma>\sigma^*$. Consider the reflection of $\Delta_3(\sigma)$ about the σ-axis; i.e., $\hat{\Delta}_3(\sigma)=-\Delta_3(\sigma)$. Then $\hat{\Delta}_3(\sigma)>0$ if $\sigma<\sigma^*$ and $\hat{\Delta}_3(\sigma)<0$ if $\sigma>\sigma^*$. Next, choose $\hat{\sigma}\in(\sigma_1,\sigma_2)$ and shift the function $\hat{\Delta}_3(\sigma)$ to the left as follows

$$
\begin{aligned}
\tilde{\Delta}_3(\sigma) &= \hat{\Delta}_3\left(\sigma+(\sigma^*-\hat{\sigma})\right) = -\Delta_3\left(\sigma+(\sigma^*-\hat{\sigma})\right)\\
&= U(\mathbf{z_2},\mathbf{t_2}+\sigma^*-\hat{\sigma}+\sigma)-U(\mathbf{z_1},\mathbf{t_1}+\sigma^*-\hat{\sigma}+\sigma),
\end{aligned}
$$

so that for the shifted function (see Figure A.3 in the Appendix):

$$\tilde{\Delta}_3(\sigma_1)>0, \tilde{\Delta}_3(\hat{\sigma})=0, \tilde{\Delta}_3(\sigma_2)<0, \text{ and } \tilde{\Delta}_3(\sigma')<0.$$

Define the mixtures $(\mathbf{w_1},\mathbf{l_1})=(\mathbf{z_2},\mathbf{t_2}+\sigma^*-\hat{\sigma})\lambda(\mathbf{x},\mathbf{t})$ and $(\mathbf{w_2},\mathbf{l_2})=(\mathbf{z_1},\mathbf{t_1}+\sigma^*-\hat{\sigma})\lambda(\mathbf{y},\mathbf{s})$. Analyse the function

$$
\begin{aligned}
\Delta_4(\sigma) &= U(\mathbf{w_1},\mathbf{l_1}+\sigma)-U(\mathbf{w_2},\mathbf{l_2}+\sigma)\\
&= \lambda\left(U(\mathbf{z_2},\mathbf{t_2}+\sigma^*-\hat{\sigma}+\sigma)-U(\mathbf{z_1},\mathbf{t_1}+\sigma^*-\hat{\sigma}+\sigma)\right)\\
&\quad +(1-\lambda)\left(U(\mathbf{x},\mathbf{t}+\sigma)-U(\mathbf{y},\mathbf{s}+\sigma)\right)\\
&= \lambda\tilde{\Delta}_3(\sigma)+(1-\lambda)\Delta_1(\sigma).
\end{aligned}
$$

Choosing λ sufficiently small we obtain (see Figure A.4 in the Appendix):

$$\Delta_4(\sigma_1)>0, \Delta_4(\sigma_2)<0, \text{ and } \Delta_4(\sigma')>0.$$

Therefore, we have arrived at a contradiction to the one-switch property.

Case 2. We next assume that there exist $\sigma_1^*<\sigma_2^*$ such that $\Delta_3(\sigma)=0$ if and only if $\sigma\in[\sigma_1^*,\sigma_2^*]\subset(\sigma_2,\sigma')$. Also, by the one-switch property, we must have $\Delta_3(\sigma)<0$ if $\sigma<\sigma_1^*$, and $\Delta_3(\sigma)>0$ if $\sigma>\sigma_2^*$. We consider the reflection of $\Delta_3(\sigma)$ about the σ-axis

[8]Note that for all figures in Appendix only sign is relevant in the vertical dimension.

and shift it to the left by a small amount $\epsilon > 0$ to obtain $\bar{\Delta}_3(\sigma)$ as follows

$$\bar{\Delta}_3(\sigma) \;=\; -\Delta_3(\sigma+\varepsilon) \;=\; U(\mathbf{z_2},\mathbf{t_2}+\varepsilon+\sigma) - U(\mathbf{z_1},\mathbf{t_1}+\varepsilon+\sigma).$$

Then (see Figure A.5 in the Appendix):

$$\bar{\Delta}_3(\sigma) > 0 \text{ if } \sigma < \sigma_1^* - \varepsilon, \bar{\Delta}_3(\sigma) = 0 \text{ if } \sigma \in [\sigma_1^* - \varepsilon, \sigma_2^* - \varepsilon], \text{ and } \bar{\Delta}_3(\sigma) < 0 \text{ if } \sigma > \sigma_2^* - \varepsilon.$$

Define the mixtures $(\mathbf{v_1},\mathbf{k_1}) = (\mathbf{z_2},\mathbf{t_2}+\varepsilon)\lambda(\mathbf{z_1},\mathbf{t_1})$ and $(\mathbf{v_2},\mathbf{k_2}) = (\mathbf{z_1},\mathbf{t_1}+\varepsilon)\lambda(\mathbf{z_2},\mathbf{t_2})$. Analyse the function

$$\begin{aligned}
\bar{\Delta}_4(\sigma) &= U(\mathbf{v_1},\mathbf{k_1}+\sigma) - U(\mathbf{v_2},\mathbf{k_2}+\sigma) \\
&= \lambda\left(U(\mathbf{z_2},\mathbf{t_2}+\varepsilon+\sigma) - U(\mathbf{z_1},\mathbf{t_1}+\varepsilon+\sigma)\right) + \\
&\quad (1-\lambda)\left(U(\mathbf{z_1},\mathbf{t_1}+\sigma) - U(\mathbf{z_2},\mathbf{t_2}+\sigma)\right) \\
&= \lambda\bar{\Delta}_3(\sigma) + (1-\lambda)\Delta_3(\sigma).
\end{aligned}$$

Choosing λ sufficiently small we obtain (see Figure A.6 in the Appendix):

$$\bar{\Delta}_4(\sigma_1) < 0, \bar{\Delta}_4(\sigma_1^*) = 0, \text{ and } \bar{\Delta}_4(\sigma_2^*) < 0.$$

Finally, mixing $(b,p)\lambda(\mathbf{v_1},\mathbf{k_1})$ and $(a,q)\lambda(\mathbf{v_2},\mathbf{k_2})$ we have

$$\begin{aligned}
\Delta_5(\sigma) &= \lambda\left(U(b,p+\sigma) - U(a,q+\sigma)\right) + (1-\lambda)\left(U(\mathbf{v_1},\mathbf{k_1}+\sigma) - U(\mathbf{v_2},\mathbf{k_2}+\sigma)\right) \\
&= -\lambda\Delta_2(\sigma) + (1-\lambda)\bar{\Delta}_4(\sigma).
\end{aligned}$$

Recall that $\Delta_2(\sigma) < 0$ for all σ. Letting λ be sufficiently small we obtain a double switch (see Figure A.7 in the Appendix):

$$\Delta_5(\sigma_1) < 0, \Delta_5(\sigma_1^*) > 0, \text{ and } \Delta_5(\sigma_2^*) < 0,$$

which is a contradiction. □

The implication of Proposition 4.20 is as follows:

Corollary 4.21. *If preferences $\succcurlyeq$ on $\mathcal{A}$ have a DEU representation, then the one-switch property is equivalent to the weak one-switch property.*

4.3.2 The weak one-switch property for dated outcomes and impatience

In this section we consider preferences $\succcurlyeq$ on $\mathcal{A}_1$ and return to the initial assumption that $X = [0, \infty)$.

Definition 4.22. *We say that* $\succcurlyeq$ *exhibits the* weak one-switch property for dated outcomes *if* $\succcurlyeq$ *exhibits the weak one-switch property on* $\mathcal{A}_1$.

The following proposition gives a partial characterisation of the weak one-switch property for dated outcomes.

Proposition 4.23. *Let* $\succcurlyeq$ *restricted to* $\mathcal{A}_1$ *has a DU representation. Then preferences* $\succcurlyeq$ *exhibit the weak one-switch property for dated outcomes if* $\succcurlyeq$ *exhibit DI or II.*

Proof. We show that if $\succcurlyeq$ exhibits II or DI it must also exhibit the weak one-switch property for dated outcomes. The proof is by contradiction. Suppose that for some $(x, t), (y, s) \in \mathcal{A}_1$ and some σ, ε such that $0 < \sigma < \varepsilon$ we can have:

$$(x, t) \succ (y, s), \tag{4.32}$$

$$(x, t + \sigma) \prec (y, s + \sigma), \tag{4.33}$$

$$(x, t + \varepsilon) \succ (y, s + \varepsilon). \tag{4.34}$$

W.l.o.g. assume that $t < s$. Then we also must have that $x < y$, otherwise (4.33) contradicts impatience and monotonicity. By continuity we can find $s' \in (t, s)$ such that $(x, t) \sim (y, s')$. Therefore, by II we have $(x, t + \sigma) \succcurlyeq (y, s' + \sigma)$. By impatience, $(y, s' + \sigma) \succ (y, s + \sigma)$, hence, by transitivity, $(x, t + \sigma) \succ (y, s + \sigma)$, which is a contradiction to (4.33). Therefore, we have shown that if $\succcurlyeq$ does not satisfy the one-switch property for dated outcomes it also does not exhibit II.

The proof for DI is analogous. Indeed, consider $(x, t + \sigma) \prec (y, s + \sigma)$, with $t < s$, $x < y$, as before. By continuity we can find $s'' > s$ such that $(x, t + \sigma) \sim (y, s'' + \sigma)$. Since $\varepsilon > \sigma$, let $\varepsilon = \sigma + \gamma$ for some $\gamma > 0$. From the equivalence $(x, t + \sigma) \sim (y, s'' + \sigma)$ it follows by DI that $(x, t + \sigma + \gamma) \preccurlyeq (y, s'' + \sigma + \gamma)$, or, equivalently, $(x, t + \varepsilon) \preccurlyeq (y, s'' + \varepsilon)$. Since $s'' > s$, impatience implies that $(y, s'' + \varepsilon) \prec (y, s + \varepsilon)$. Therefore, we must have $(x, t + \varepsilon) \prec (y, s + \varepsilon)$, which brings us to a contradiction to (4.34). □

We have not been able to establish the converse to Proposition 4.23. The arguments used in Lemma 4.11 do not adapt straightforwardly to the present situation. However, Proposition 4.23, together with previous characterisation of the one-switch property for dated outcomes (Lemma 4.11), already imply that the one-switch property for dated outcomes and the weak one-switch property for dated outcomes are not equivalent for intertemporal preferences with a DU representation.

4.4 Discussion

First of all, this chapter fills a gap in Bell's original characterisation [12, Proposition 8] of discount functions compatible with the one-switch property for sequences of dated outcomes. We showed that functions of the linear times exponential form also have this property and that such discount functions exhibit strictly increasing impatience. Although decreasing impatience is commonly found in experiments [33], there is also much evidence for increasing impatience (see, for example, [11], [77]). To the best of our knowledge, the linear times exponential function has not been used in the literature on time preferences before. Therefore, we introduce a new type of a discount function that accommodates strictly increasing impatience and the one-switch property.

With regard to sums of exponentials, there has recently been some interest in this type of a discount function. In their experiments, McClure et al. [54] used magnetic neuro-images of individuals' brains to study intertemporal choice. They found that making a decision is a result of the interaction of two separate neural systems with different levels of impatience. To describe the involvement of these two brain areas in discounting, they suggested sum of exponential discount functions which they refer to as *double exponential* discounting. The recent empirical study by Cavagnaro et al. [20] demonstrates that double exponential discounting provides a better fit to actual time preferences than exponential, quasi-hyperbolic, proportional hyperbolic and generalized hyperbolic discounting.

A second contribution of this chapter is to clarify the definition of the original one-switch rule introduced by Bell [12], by considering two definitions: the weak one-switch property and the (strong) one-switch property. We demonstrate that if X is a mixture set and if preferences have a DEU representation, then the one-switch property is equivalent to the weak one-switch property.

Third, we prove that if preferences have a DU representation, then preferences exhibit the one-switch property for dated outcomes if and only if they exhibit either strictly increasing impatience, or strictly decreasing impatience, or constant impatience. A partial analogue is obtained for the weak one-switch property for dated outcomes. That is, the preferences exhibit the weak one-switch property for dated outcomes if they exhibit increasing impatience or decreasing impatience.

5

Discussion and outlook

This thesis investigated three topics in intertemporal choice: axiomatization of time preferences, aggregation of time preferences with decreasing impatience and the effect of delay on preferences for sequences of outcomes. We used various time preference frameworks in this thesis with preferences being expressed over consumption streams (the discrete time setting) or single dated outcomes and sequences of dated outcomes (in the continuous time setting). We also considered deterministic and stochastic outcomes.

In Chapter 2 we demonstrated that if we adopt the Anscombe and Aumann [8] framework, adapted to an intertemporal setting, then we can get simple specializations of the AA representation to exponential and quasi-hyperbolic discounting in finite and infinite horizons. To obtain these representations separability must be assumed.[1] The mixture independence axiom, while a strong assumption, gives a nice short-cut to the desired separability and allows us to replace a set of independence conditions of different types with a single axiom. As a result, application of the AA representation theorem gave us a simple axiomatization for exponential and quasi-hyperbolic discounting with relatively straightforward proofs. This is an advantage of our axiomatization system compared to [44] and [58]. It is worth mentioning that the AA set-up was previously used in the time preferences context by Wakai [79] to axiomatize preferences for spread-

[1]Note that the assumption of separability is considered to be controversial as it is empirically questionable [64], [80], [27].

ing bad and good outcomes evenly over time (which is a completely different type of preferences).

The results of Chapter 3 extend our knowledge of discounting behaviour in a group setting or when there is some uncertainty about the appropriate discount function. The set-up of Chapter 3 considers preferences over single dated outcomes rather than over consumption streams. This chapter highlighted and clarified the existing inconsistency in the time preference literature on the terminology of decreasing impatience and present bias. This study analysed convenient instruments (time preference rate, index of DI) introduced by Prelec [63] for characterising different types of impatience (DI/II, strictly DI/ strictly II) and comparing them among different decision-makers. The key contribution of Chapter 3 is a generalization of the result of Prelec [63] and Jackson and Yariv [46] on the aggregation of discount functions. For discount functions from the same DI class, we proved that the aggregated discount function is strictly more DI than each of the individual discount functions. Under additional smoothness assumptions, we study mixtures of arbitrary discount functions. Using Prelec's [63] Index of DI as a local measure of DI at each point of time, we demonstrated that the index of the mixed discount function is always weakly larger than the minimum of the indices of the component functions. Another contribution of this chapter is an analogue of Weitzman's result [81] on the aggregation of exponential discount functions for the case of hyperbolic discounting. We showed that the hyperbolic result is very different to the exponential case. The long term certainty-equivalent hyperbolic discount rate is the weighted harmonic mean of the possible hyperbolic discount rates.

Chapter 4 enhances our understanding of the effect of attitude to delay on the shape of the discount function. More precisely, we study the one-switch property, which allows the preference ranking between two sequences of dated outcomes to switch at most once as delay increases. Recall from Chapter 3 that the sum of exponential discount function exhibits strictly DI. The research question of Chapter 4 was motivated by what can be considered as an inverse problem: what property of time preferences (without assuming the aggregation this time) implies that the discount function should be a sum of exponentials? In fact, such a characterisation already existed in the work of Bell [12], though this seems not to be widely known. Bell [12] mainly analyses risk preferences, where the one-switch property refers to the effect of wealth on lottery preference, but he also has a lesser known section that adapts his results to the intertemporal context. Given the recent interest in the experimental literature [20, 54] in the sum of exponential discount functions, Bell's result suggests that the one-switch property may provide useful motivation for this type of discounting. However, in Chapter 4 we show that the one-switch property is also compatible with another form of discount function: the linear times exponential. To the best of our knowledge, this type of discount function has not been previously used in intertemporal context. As demonstrated in the chapter, linear times exponential discount functions exhibit strictly II. While strictly II has not been a very frequent experimental observation [33], some recent empirical findings support this type of impatience [77, 74, 16]. We also analysed the distinction between the weak one-switch property and the (strong) one-switch property in the time preference context. While this distinction is inconsequential in the risk set-up with an expected utility representation, matters are not so clear for the intertemporal context, even assuming a discounted utility representation. We showed, however, that these two properties are equivalent in a set-up analogous to that of Anscombe and Aumann [8]. This is a continuous-time version of the set-up used in Chapter 2.

In terms of recommendations for further research work, Chapter 4 has thrown up a few questions in need of investigation.

Identifying axiomatic foundations for discounted expected utility (DEU) is one of the most interesting open questions. We are not familiar with any axiomatic system justifying this type of representation of time preferences in a continuous time environment. Another natural progression of this work would be to examine the role of separability and mixture linearity in the results. Another open question is whether the weak and strong one-switch properties are equivalent beyond the AA setting. A related question is how to characterise the weak one-switch property for dated outcomes in terms of impatience.

We have only been able to provide a partial answer to the question posed above: what preference property implies sum of exponential discounting? Our results imply that the one-switch property and DI suffice, but are not necessary. Finding the necessary conditions for sum of exponential discounting can be another direction of future work.

It would be interesting to characterise a different one-switch property: one in which the switch occurs between increasing and decreasing impatience as delay increases (see [77] for experimental evidence on this behaviour). It seems natural to assume that relaxing the one-switch property for sequences of dated outcomes to the double-switch property might accommodate such behaviour. Further research needs to be carried out to explore this idea and define the representation.

The one-switch property seems attractive from theoretical point of view, so it is of some interest to test it empirically in the time preference context. The experiments conducted by Olea and Strzalecki [58] are close to testing the one-switch property but not exactly the same. The difference is that in their experimental set-up an outcome at the first period always remains fixed while the rest of the sequence of dated outcomes is delayed. Designing and carrying out experiments to study the effect of delay on the ranking of sequences of dated outcomes might be a fruitful area for further work.

There are also a few avenues for further research based on other chapters. One possible extension of the work in Chapter 2 is axiomatization of generalized hyperbolic discounting in the AA set-up. An obvious extension of Chapter 3 is the analysis of certainty equivalents for mixtures of discount functions with two parameters, such as generalized hyperbolic discounting or quasi-hyperbolic discounting.

Appendix A: Illustrations for the proof of Proposition 4.20

In this appendix we provide the illustrations to accompany the proof of Proposition 4.20. Note that for all the figures in the appendix only the *sign* (but not the value) of a vertical coordinate of a point is relevant.

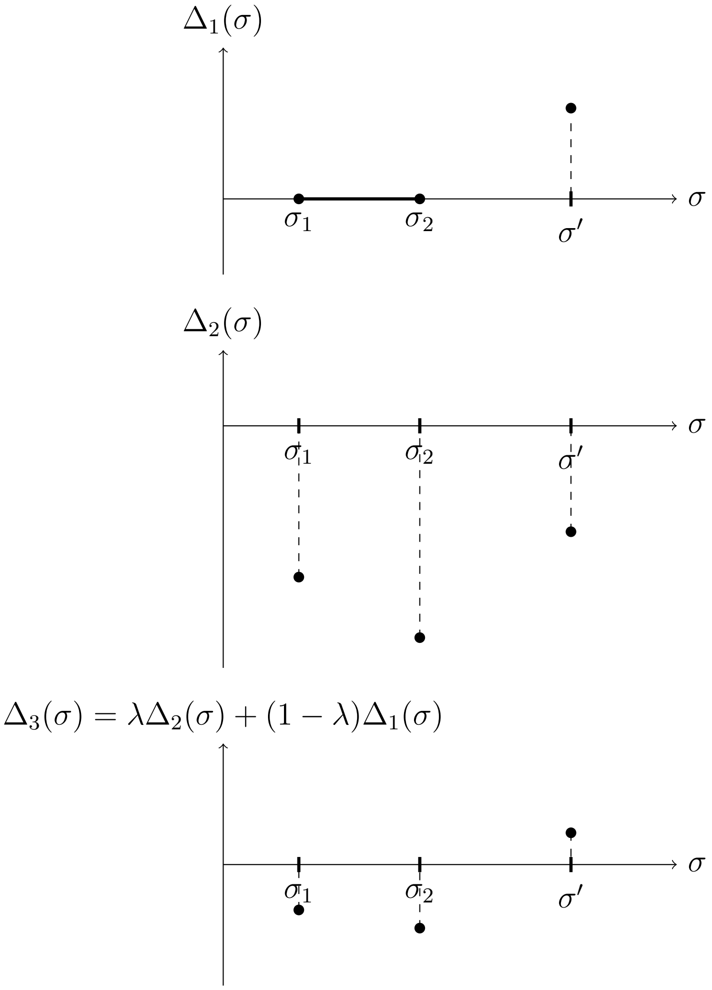

Figure A.1: The mixture of $\Delta_2(\sigma)$ and $\Delta_1(\sigma)$ (with small λ)

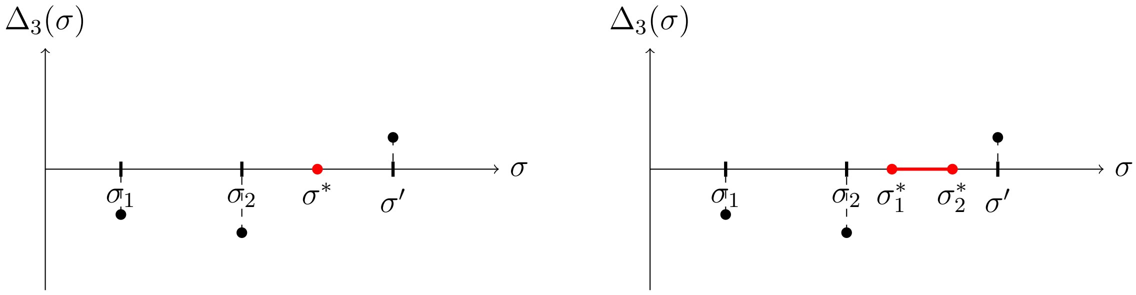

Figure A.2: Function $\Delta_3(\sigma)$ equals zero at one point or at the interval

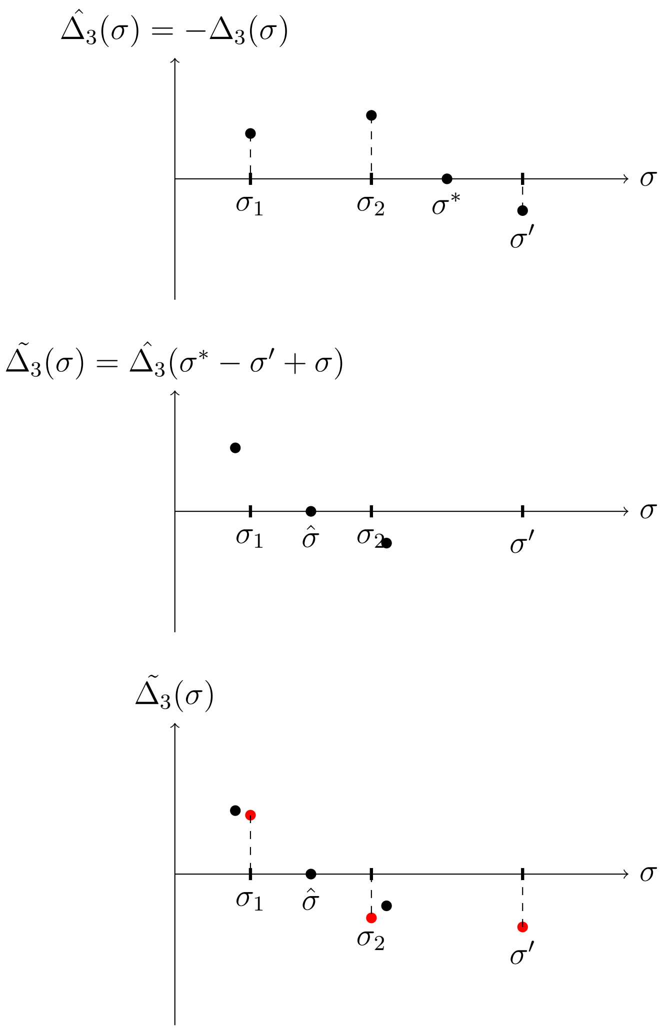

Figure A.3: Transformation of $\Delta_3(\sigma)$ to $\tilde{\Delta_3}(\sigma)$, and then with the new values at indicated $\sigma_1, \sigma_2, \sigma'$

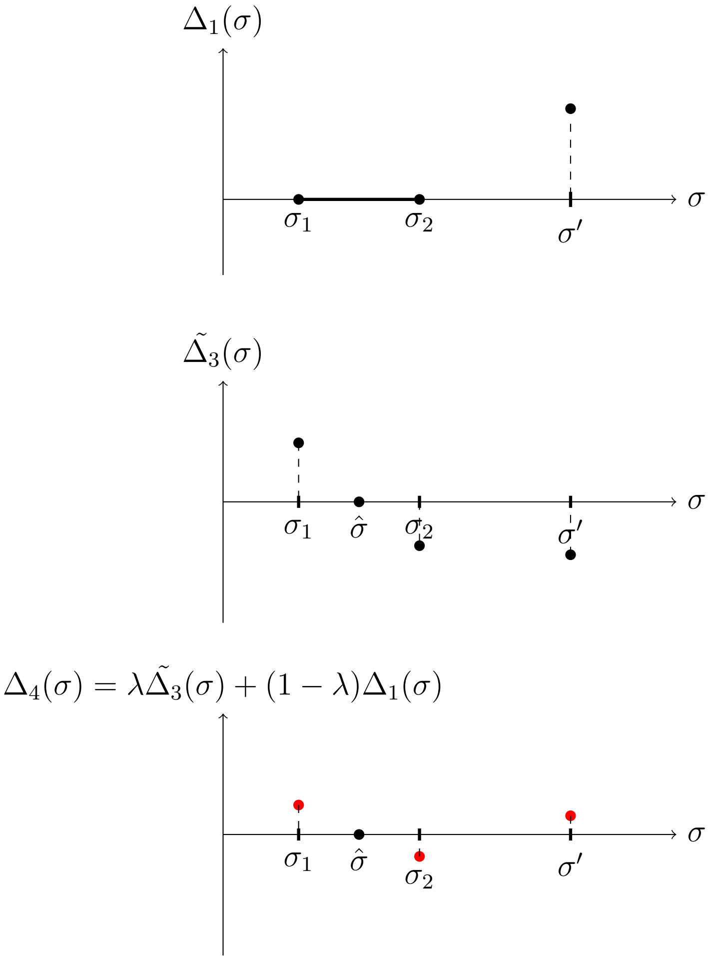

Figure A.4: The mixture of $\tilde{\Delta}_3(\sigma)$ and $\Delta_1(\sigma)$ (with small λ) produces a double switch

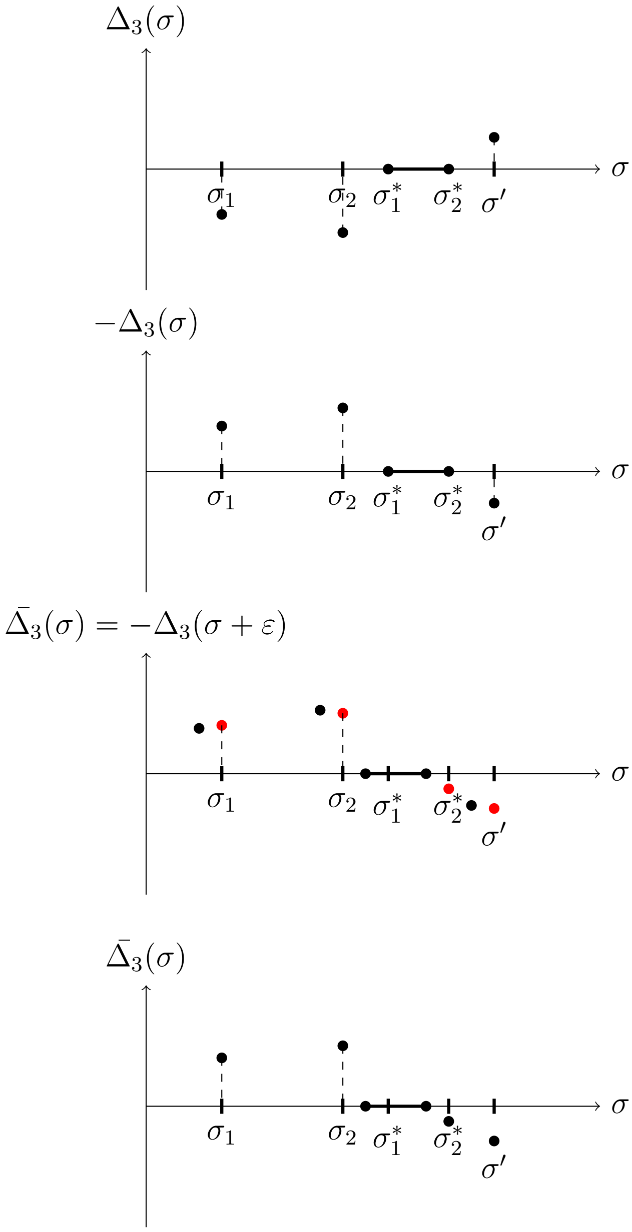

Figure A.5: Transformation of $\Delta_3(\sigma)$ to $\bar{\Delta}_3(\sigma)$ with the new values at $\sigma_1, \sigma_2, \sigma_2^*, \sigma'$

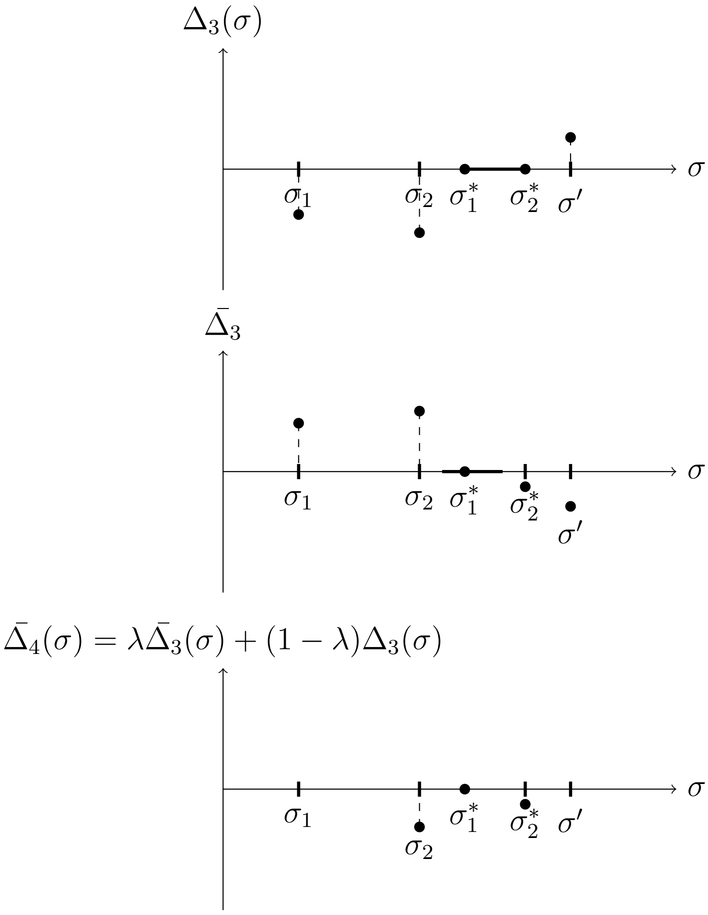

Figure A.6: The mixture of $\bar{\Delta}_3(\sigma)$ and $\Delta_3(\sigma)$ (with small λ) at the three points $\sigma_2, \sigma_1^*, \sigma_2^*$

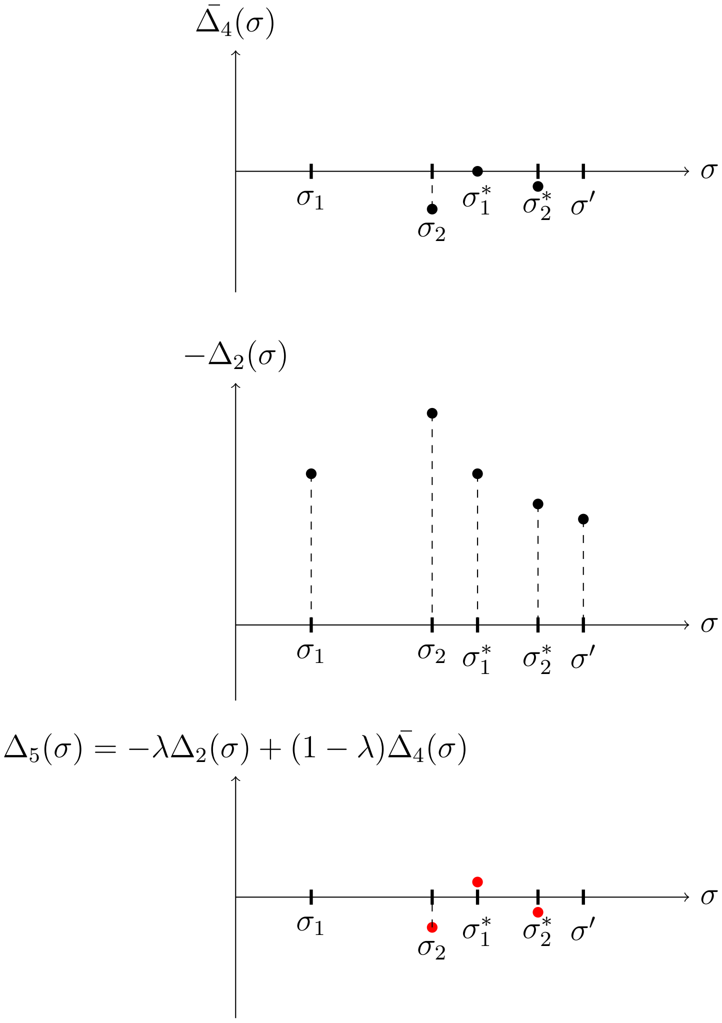

Figure A.7: The mixture of $-\Delta_2(\sigma)$ and $\bar{\Delta}_4(\sigma)$ (with small λ) produces a double switch. It is sufficient to consider the value of $\Delta_5(\sigma)$ at the three points $\sigma_2, \sigma_1^*, \sigma_2^*$

Bibliography

[1] A. E. Abbas and D. E. Bell. One-switch independence for multiattribute utility functions. *Operations Research*, 59(3):764–771, 2011.

[2] A. E. Abbas and D. E. Bell. Methods-one-switch conditions for multiattribute utility functions. *Operations Research*, 60(5):1199–1212, 2012.

[3] A. E. Abbas and D. E. Bell. Ordinal one-switch utility functions. *Operations Research*, 63(6):1411–1419, 2015.

[4] A. Al-Nowaihi and S. Dhami. A note on the Loewenstein-Prelec theory of intertemporal choice. *Mathematical Social Sciences*, 52(1):99–108, 2006.

[5] J. M. Almira. On Popoviciu-Ionescu functional equation. *Annales Mathematicae Silesianae*, 30(1):5–15, 2016.

[6] N. Anchugina. A simple framework for the axiomatization of exponential and quasi-hyperbolic discounting. *Theory and Decision*, 2016. doi:10.1007/s11238-016-9566-8.

[7] N. Anchugina, M. Ryan, and A. Slinko. Aggregating time preferences with decreasing impatience. *arXiv preprint arXiv:1604.01819*, 2016.

[8] F. J. Anscombe and R. J. Aumann. A definition of subjective probability. *Annals of Mathematical Statistics*, 34(1):199–205, 1963.

[9] K. J. Arrow. *Aspects of the theory of risk-bearing.* Helsinki: Yrjö Jahnssonin Säätiö, 1965.

[10] E. Artin. *The gamma function.* New York: Holt, Rinehart and Winston, 1964.

[11] A. E. Attema, H. Bleichrodt, K. I. M. Rohde, and P. P. Wakker. Time-tradeoff sequences for analyzing discounting and time inconsistency. *Management Science*, 56(11):2015–2030, 2010.

[12] D. E. Bell. One-switch utility functions and a measure of risk. *Management Science*, 34(12):1416–1424, 1988.

[13] D. E. Bell. Risk, return, and utility. *Management Science*, 41(1):23–30, 1995.

[14] D. E. Bell and P. C. Fishburn. Utility functions for wealth. *Journal of Risk and Uncertainty*, 20(1):5–44, 2000.

[15] D. E. Bell and P. C. Fishburn. Strong one-switch utility. *Management Science*, 47(4):601–604, 2001.

[16] H. Bleichrodt, Y. Gao, and K. I. M. Rohde. A measurement of decreasing impatience for health and money. *Journal of Risk and Uncertainty*, 52(3):213–231, 2016.

[17] H. Bleichrodt, K. I. M. Rohde, and P. P. Wakker. Koopmans' constant discounting for intertemporal choice: a simplification and a generalization. *Journal of Mathematical Psychology*, 52(6):341–347, 2008.

[18] H. Bleichrodt, K. I. M. Rohde, and P. P. Wakker. Non-hyperbolic time inconsistency. *Games and Economic Behavior*, 66(1):27–38, 2009.

[19] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge: Cambridge University Press, 2004.

[20] D. R. Cavagnaro, G. J. Aranovich, S. M. McClure, M. A. Pitt, and J. I. Myung. On the functional form of temporal discounting: an optimized adaptive test. *Journal of Risk and Uncertainty*, 52(3):233–254, 2016.

[21] R. Chark, S. H. Chew, and S. Zhong. Extended present bias: a direct experimental test. *Theory and Decision*, 79(1):151–165, 2015.

[22] Z. Cvetkovski. *Inequalities: theorems, techniques and selected problems.* Heidelberg: Springer, 2012.

[23] G. Debreu. Topological methods in cardinal utility theory. In K.J. Arrow, S. Karlin, and P. Suppes, editors, *Mathematical methods in the social sciences*, pages 16–26. Stanford, CA: Stanford University Press, 1960.

[24] J. R. Doyle. Survey of time preference, delay discounting models. *Judgment and Decision Making*, 8(2):116–135, 2013.

[25] J. E. J. Ebert and D. Prelec. The fragility of time: time-insensitivity and valuation of the near and far future. *Management Science*, 53(9):1423–1438, 2007.

[26] L. G. Epstein. Stationary cardinal utility and optimal growth under uncertainty. *Journal of Economic Theory*, 31(1):133–152, 1983.

[27] K. M. Ericson and J. Noor. Delay functions as the foundation of time preference: Testing for separable discounted utility. Technical Report 21095, National Bureau of Economic Research, 2015.

[28] J. D. Farmer and J. Geanakoplos. Hyperbolic discounting is rational: Valuing the far future with uncertain discount rates. Technical Report 1719, Cowles Foundation for Research in Economics, Yale University, 2009.

[29] P. Fishburn and W. Edwards. Discount-neutral utility models for denumerable time streams. *Theory and Decision*, 43(2):139–166, 1997.

[30] P. C. Fishburn. *Utility theory for decision making*. New York: Wiley, 1970.

[31] P. C. Fishburn. *The foundations of expected utility*, volume 31. Dordrecht: Reidel Publishing Co, 1982.

[32] P. C. Fishburn and A. Rubinstein. Time preference. *International Economic Review*, 23(3):677–694, 1982.

[33] S. Frederick, G. Loewenstein, and T. O'Donoghue. Time discounting and time preference: a critical review. *Journal of Economic Literature*, 40(2):351–401, 2002.

[34] J. Geanakoplos, W. Sudderth, and O. Zeitouini. Asymptotic behavior of a stochastic discount rate. *Indian Journal of Statistics*, 76-A(1):150–157, 2014.

[35] I. Gilboa and D. Schmeidler. Maxmin expected utility with non-unique prior. *Journal of Mathematical Economics*, 18(2):141–153, 1989.

[36] J.-M. Grandmont. Continuity properties of a von Neumann-Morgenstern utility. *Journal of Economic Theory*, 4(1):45–57, 1972.

[37] S. Grant and T. Van Zandt. Expected utility theory. In P. Anand, P. Pattanaik, and C. Puppe, editors, *The handbook of rational and social choice*, pages 21–68. Oxford: Oxford University Press, 2009.

[38] F. Gul. Savage's theorem with a finite number of states. *Journal of Economic Theory*, 57(1):99–110, 1992.

[39] J. Gurland and J. Sethuraman. Shorter communication: reversal of increasing failure rates when pooling failure data. *Technometrics*, 36(4):416–418, 1994.

[40] J. Gurland and J. Sethuraman. How pooling failure data may reverse increasing failure rates. *Journal of the American Statistical Association*, 90(432):1416–1423, 1995.

[41] Y. Halevy. Strotz meets Allais: diminishing impatience and the certainty effect. *American Economic Review*, 98(3):1145–1162, 2008.

[42] C. M. Harvey. Value functions for infinite-period planning. *Management Science*, 32(9):1123–1139, 1986.

[43] C. M. Harvey. Proportional discounting of future costs and benefits. *Mathematics of Operations Research*, 20(2):381–399, 1995.

[44] T. Hayashi. Quasi-stationary cardinal utility and present bias. *Journal of Economic Theory*, 112(2):343–352, 2003.

[45] R. J. Herrnstein. Self-control as response strength. In C.M. Bradshaw, E. Szabadi, and C.F. Lowe, editors, *Quantification of steady-state operant behavior*, pages 3–20. Amsterdam: Elsevier/North Holland Biomedical Press, 1981.

[46] M. O. Jackson and L. Yariv. Present bias and collective dynamic choice in the lab. *American Economic Review*, 104(12):4184–4204, 2014.

[47] M. O. Jackson and L. Yariv. Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–178, 2015.

[48] D. T. Jamison and J. Jamison. Characterizing the amount and speed of discounting procedures. *Journal of Benefit-Cost Analysis*, 2(2):1–56, 2011.

[49] T. C. Koopmans. Stationary ordinal utility and impatience. *Econometrica*, 28(2):287–309, 1960.

[50] T. C. Koopmans. Representation of preference orderings over time. In C.B. McGuire and R. Radner, editors, *Decision and organization*, pages 79–100. Amsterdam: North Holland, 1972.

[51] T. C. Koopmans, P. A. Diamond, and R. E. Williamson. Stationary utility and time perspective. *Econometrica*, 32(1/2):82–100, 1964.

[52] D. Laibson. Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics*, 112(2):443–478, 1997.

[53] G. Loewenstein and D. Prelec. Anomalies in intertemporal choice: evidence and an interpretation. *Quarterly Journal of Economics*, 107(2):573–597, 1992.

[54] S. M. McClure, K. M. Ericson, D. I. Laibson, G. Loewenstein, and J. D. Cohen. Time discounting for primary rewards. *The Journal of Neuroscience*, 27(21):5796–5804, 2007.

[55] R. Mickens. *Difference equations.* New York: Van Nostrand Reinhold, 1991.

[56] J. Noor. Decreasing impatience and the magnitude effect jointly contradict exponential discounting. *Journal of Economic Theory*, 144(2):869–875, 2009.

[57] E. A. Ok and Y. Masatlioglu. A theory of (relative) discounting. *Journal of Economic Theory*, 137(1):214–245, 2007.

[58] J. L. M. Olea and T. Strzalecki. Axiomatization and measurement of quasi-hyperbolic discounting. *Quarterly Journal of Economics*, 129(3):1449–1499, 2014.

[59] J. Pan, C. S. Webb, and H. Zank. An extension of quasi-hyperbolic discounting to continuous time. *Games and Economic Behavior*, 89:43–55, 2015.

[60] E. S. Phelps and R. A. Pollak. On second-best national saving and game-equilibrium growth. *Review of Economic Studies*, 35(2):185–199, 1968.

[61] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964.

[62] J. W. Pratt. Aversion to one risk in the presence of others. *Journal of Risk and Uncertainty*, 1(4):395–413, 1988.

[63] D. Prelec. Decreasing impatience: a criterion for non-stationary time preference and hyperbolic discounting. *Scandinavian Journal of Economics*, 106(3):511–532, 2004.

[64] D. Prelec and G. Loewenstein. Decision making over time and under uncertainty: a common approach. *Management Science*, 37(7):770–786, 1991.

[65] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3):373–383, 1963.

[66] J. Quiggin and J. Horowitz. Time and risk. *Journal of Risk and Uncertainty*, 10(1):37–55, 1995.

[67] F. Radó. Caractérisation de l'ensemble des intégrales des équations différentielles linéaires homogènes à coefficients constants d'ordre donné. *Mathematica (Cluj)*, 4(27):131–143, 1962.

[68] D. Read. Is time-discounting hyperbolic or subadditive? *Journal of Risk and Uncertainty*, 23(1):5–32, 2001.

[69] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *A Series of Comprehensive Studies in Mathematics*. Heidelberg: Springer-Verlag, 1998.

[70] K. I. M. Rohde. The hyperbolic factor: a measure of time inconsistency. *Journal of Risk and Uncertainty*, 41(2):125–140, 2010.

[71] M. J. Ryan. Generalizations of SEU: a geometric tour of some non-standard models. *Oxford Economic Papers*, 61(2):327–354, 2009.

[72] P. A. Samuelson. A note on measurement of utility. *Review of Economic Studies*, 4(2):155–161, 1937.

[73] L. J. Savage. *The foundations of statistics*. New York: John Wiley and Sons, 1972.

[74] S. Sayman and A. Öncüler. An investigation of time inconsistency. *Management Science*, 55(3):470–482, 2009.

[75] P. D. Sozou. On hyperbolic discounting and uncertain hazard rates. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 265(1409):2015–2020, 1998.

[76] O. Stein. Twice differentiable characterizations of convexity notions for functions on full dimensional convex sets. *Schedae Informaticae*, 21:55–63, 2012.

[77] K. Takeuchi. Non-parametric test of time consistency: present bias and future bias. *Games and Economic Behavior*, 71(2):456–478, 2011.

[78] J. Von Neumann and O. Morgenstern. *Theory of games and economic behavior*. Princeton: Princeton University Press, 1947.

[79] K. Wakai. A model of utility smoothing. *Econometrica*, 76(1):137–153, 2008.

[80] L. Wathieu. Habits and the anomalies in intertemporal choice. *Management Science*, 43(11):1552–1563, 1997.

[81] M. L. Weitzman. Why the far-distant future should be discounted at its lowest possible rate. *Journal of Environmental Economics and Management*, 36(3):201–208, 1998.

[82] L. Xue. The bargaining within. *Economics Letters*, 101(2):145–147, 2008.

[83] E. R. Young. Generalized quasi-geometric discounting. *Economics Letters*, 96(3):343–350, 2007.

[84] S. Zuber. The aggregation of preferences: can we ignore the past? *Theory and Decision*, 70(3):367–384, 2011.