Aggregating time preferences with decreasing impatience

Nina Anchugina[1], Matthew Ryan[2], and Arkadii Slinko[3]

[1]Centre for Social Data Analytics, Auckland University of Technology
[2]School of Economics, Auckland University of Technology
[3]Department of Mathematics, University of Auckland
nina.anchugina@aut.ac.nz, mryan@aut.ac.nz, a.slinko@auckland.ac.nz

July 2017

Abstract. It is well-known that for a group of time-consistent decision makers their collective time preferences may become time-inconsistent. In particular, Jackson and Yariv ([15]) show that aggregating heterogeneous exponential discount functions yields a function that exhibits what they call "present bias". In a continuous-time setting "present bias" is equivalent to strictly decreasing impatience (DI). Applying the notion of comparative DI introduced by Prelec ([22]), we generalise Jackson and Yariv's result: mixing any finite number of heterogeneous discount functions from the same DI equivalence class (such as exponential discount functions) yields a mixture that exhibits uniformly more DI than each component. We also obtain necessary and sufficient conditions for mixtures of DI-ordered – but not necessarily DI-equivalent – functions to be uniformly more DI than the least DI component. For suitably smooth discount functions, Prelec ([22]) also defines a local index of DI. We use this to obtain local versions of our results, which apply to mixtures of any finite set of (smooth) discount functions. Our main theorems generalise a number of specialised results in the decision theory and survival analysis literatures.

Keywords: Hyperbolic discounting, Present bias, Decreasing impatience, Aggregation.
JEL Classification: D71, D90.

1 Introduction

In the context of intertemporal choice, there are various situations which call for the use of *mixtures* (weighted averages) of discount functions to evaluate streams of outcomes. Consider, for example, a utilitarian Social Planner aggregating the preferences of a group of discounted utility maximisers over a common consumption stream. If all members of the group share a common utility function but have heterogeneous discount functions, the utilitarian will maximise discounted utility using a discount function that is a mixture of the individual discount functions. Jackson and Yariv ([16]) further show that if each individual discount function is exponential, any social welfare function that respects unanimity *must* be utilitarian. Mixtures of discount functions also arise when a Planner is uncertain about the appropriate discount factor to apply to distant future consumption streams, as in the analysis of Gollier and Weitzmann ([28], [12]). Recent evidence from neuropsychology even suggests that individuals make intertemporal choices by aggregating different internal motives, leading to preference representations involving mixtures of discount functions ([18]).

It is therefore of interest to understand the properties of mixed discount functions. It is well known that these properties may differ in significant ways from those of the functions being mixed. Jackson and Yariv demonstrate ([15, Proposition 1]) that any non-trivial mixture of heterogeneous exponential discount functions will exhibit strictly decreasing impatience – whenever two dated outcomes have the same discounted utility (with respect to the mixed discount function), delaying both dated outcomes by the same amount of time will produce a strict preference for the later of the two.[1] Similar phenomena have been observed by several other authors, including, most recently, by Pennesi ([20, Theorem 1]).

Related results may also be found in reliability theory, or survival analysis, where mixtures of survival functions are considered.[2] A survival function is the decumulative distribution function for a random "failure" time. Such functions therefore have similar mathematical properties to discount functions. Indeed, discounting is sometimes motivated by uncertainty about one's time of death. When pooling data on the same "component" installed in different "machines", (more generally, when considering the survival function for a population comprised of heterogeneous sub-populations), the survival function for a randomly selected component – the proportion of components expected to survive to a given time – is described by the average of the individual survival functions. Sozou ([25]) observes that a continuous mixture of survival functions with the exponential form $S(t) = \exp(-\rho t)$, in which ρ is uniformly distributed on $(0, 1)$, produces an hyperbolic function.[3] More generally, Barlow, Marshall and Proschan ([4, Theorem 3.4]) prove that any mixture of exponential survival functions will exhibit a decreasing failure rate, which is equivalent to a decreasing rate of time preference in the discounting context.[4] Gurland and Sethuraman ([13]) even provide examples of survival functions with *increasing* failure

[1] Jackson and Yariv ([15]) call this phenomenon *present bias*.

[2] See the survey by Block in [6].

[3] See also [15, Example 1].

[4] Recall that if $D : [0, \infty) \to (0, 1]$ is a differentiable discount function, then its associated rate of time

rates whose mixtures exhibit failure rates that are decreasing.

There are also well-known results on the asymptotic behaviour of failure rates for mixtures of survival functions – exponential or otherwise – showing convergence to the lowest asymptotic failure rate amongst the functions being mixed (see [7, Corollary 2.1], which refines an earlier result of Block and Joe [5]). Weitzman ([28]) and others have demonstrated similar results for mixtures of discount functions.

A common theme emerges from these various contributions. Mixing of heterogeneous discount (or survival) functions enhances the tendency for impatience (or failure rates) to diminish over time. The intuition for this general tendency is easy to grasp by considering the case of survival functions. Let $S_1, S_2, \ldots, S_n$ be survival functions for the same component installed in n different machines, and let

$$S(t) = \frac{1}{n}\sum_{i=1}^{n} S_i(t)$$

be the average of these functions. Then $S(t)$ is the probability that a randomly selected component survives to time t. Its associated failure rate is

$$r_S(t) = -\frac{S'(t)}{S(t)}.$$

It is straightforward to show that

$$r_S(t) \;=\; \sum_{i=1}^{n} \left[\frac{S_i(t)}{nS(t)}\right] r_{S_i}(t).$$

The square-bracketed term is the probability that a randomly selected component comes from machine i *conditional on the component having survived until time* t. For functions with constant failure rates (i.e., exponential survival functions) it is obvious that the conditioning event will shift probability mass towards machines with lower failure rates, and that this tendency will only increase with t.

A framework for articulating the general principle that lies behind these disparate results has been suggested by Prelec ([22]). Prelec defines a "more decreasingly impatient (DI) than" relation on discount functions,[5] derived from an underlying property of intertemporal preferences. He also characterises this comparative property for the case of preferences with a discounted utility representation over dated outcomes ([10]). He shows that the discount function D_1 exhibits more decreasing impatience than D_2 if and only if $\ln D_1$ is a convex transformation of $\ln D_2$. The analogous relation on survival functions had previously been proposed and studied by Lee ([17]). If the transformation is

preference is

$$r_D(t) = -\frac{D'(t)}{D(t)}.$$

This rate is constant if, and only if, D is exponential.

[5]Throughout we use the acronym "DI" interchangeably as a noun ("decreasing impatience") and an adjective ("decreasingly impatient"), relying on context to disambiguate.

strictly convex, we will say that D_1 is "uniformly" more DI than D_2. (This concept was not defined by Prelec but is introduced in the present paper.) Prelec ([22, Corollary 4]) establishes that the mixture of two equally DI functions will be more (in fact, uniformly more) DI than either.[6]

The present paper generalises and extends Prelec's analysis in several directions.

First, we extend Prelec's result from two to n functions: we show that mixing *any finite number* of equally decreasingly impatient functions yields a mixture that is uniformly more decreasingly impatient than each (Theorem 1). The key to this extension is the fact that the sum of log-convex functions is log-convex ([2]).

Second, we study mixtures of functions that can be weakly ordered by decreasing impatience but need not be *equally* DI. The results for this case turn out to be more nuanced than one might expect. From the previous result we know that if all functions are equally DI, their mixture will uniformly dominate (in the sense of DI) each of them. It is therefore natural to suppose that the mixture of DI-ordered functions will uniformly dominate the least DI of the functions being mixed. We show that this intuition is false (Proposition 6) and establish necessary and sufficient conditions for the mixture to uniformly dominate the least DI component (Theorem 2).

Third, under additional smoothness assumptions, we are able to study mixtures of arbitrary discount functions. These functions may or may not exhibit decreasing impatience, and may or may not be weakly ordered by comparative decreasing impatience. Smoothness allows us to use the local index of decreasing impatience introduced by Prelec ([22]). This is analogous to the Arrow-Pratt index of absolute risk aversion for expected utility. For smooth discount functions, one function is more decreasingly impatient than another if the index of the former weakly dominates the index of the latter (i.e., is weakly larger at every point in time). We show that the index of the mixed discount function always weakly dominates the *minimum* of the indices of the component functions, and we provide necessary and sufficient conditions for the index of the mixture to strictly exceed this minimum at any given point (Theorem 3).

In Section 4 we show that several results from decision theory and reliability theory follow as direct corollaries of our main theorems.

Before getting to our main results, the following section reviews the notions of absolute and comparative decreasing impatience. Appendix A.1 contains some basic facts about log-convex functions which will be needed for this discussion and the subsequent analysis.

2 Preliminaries

2.1 Preferences

Let X be the set of outcomes. We 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

[6] Prelec only asserts the weaker claim – without the "uniformly" qualifier – but his proof establishes the stronger claim asserted here.

time, where 0 corresponds to the present moment. The Cartesian product $X \times T$ will be identified with the set of dated outcomes: that is, a pair $(x, t) \in X \times T$ is a dated outcome, in which a decision-maker receives x at time t and zero at all other times. Importantly, although time is indexed continuously, x is a discrete outcome received at time t, not an instantaneous flow.

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

$$U(x, t) = D(t)u(x), \tag{1}$$

for some *instantaneous utility function* $u: X \to \mathbb{R}$ that is continuous, strictly increasing and satisfies $u(0) = 0$, and some *discount function* $D: T \to (0, 1]$ that is continuous and strictly decreasing, with $D(0) = 1$ and $\lim_{t\to\infty} D(t) = 0$. In this case we also say that (u, D) is a DU representation for $\succcurlyeq$. Fishburn and Rubinstein ([10]) provide an axiomatic foundation for a discounted utility representation.

In the rest of the paper we use $\mathcal{D}$ to denote the set of functions with the properties ascribed to discount functions in the previous paragraph. All discount functions are assumed, by definition, to lie in $\mathcal{D}$. Let $\mathcal{D}^*$ denote the subset of $\mathcal{D}$ comprising the twice continuously differentiable discount functions whose first derivative is strictly negative everywhere. The assumption that $D'(t) < 0$ for all $t \in T$ is stronger than necessary to ensure that D is strictly decreasing but will be useful to ensure that Prelec's [22] index of DI is well-defined for all t.[7]

2.2 Decreasing impatience

Since discount functions are strictly decreasing, all DU maximisers are impatient. However, a decision-maker's level of impatience – as indicated by her willingness to give up outcome x at time t in exchange for a better outcome y at some later date $t + \sigma$ – may vary with t.

Definition 1 ([22])**.** *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 indifference $(x, t) \sim (y, s)$ implies $(x, t + \sigma) \preccurlyeq [\prec] (y, s + \sigma)$.*

Jackson and Yariv ([15, p. 4190]) define an analogue of strictly DI for their discrete-time environment, which they call *present bias*.[8]

Increasing impatience (II), and its strict variant, can be defined by reversing the final preference ranking in Definition 1. We focus on DI preferences for most of the paper –

[7]For similar reasons, Pratt [21] restricts attention to "smooth" von Neumann-Morgenstern (vNM) utility functions whose first derivatives are strictly positive everywhere.

[8]Note that there is some inconsistency between Jackson and Yariv's present bias definition in their 2014 paper ([15]) and that in their 2015 paper ([16]). The definition in the latter paper is weaker.

Theorem 3 being a notable exception – since this is the most empirically relevant case ([11]).

If $\succcurlyeq$ has a discounted utility representation, then the DI condition restricts only the discount function and the nature of this restriction is well known.[9]

Proposition 1 ([14, 22])**.** *Let $\succcurlyeq$ be a preference order having discounted utility representation with discount function D. The following conditions are equivalent:*

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

We will therefore say that a discount function D is (or exhibits)[10] *(strictly) DI* if it is (strictly) log-convex on $[0, \infty)$.

If $D \in \mathcal{D}^*$ its associated *rate of time preference* is the function

$$r_D(t) = -\frac{d}{dt}\ln(D(t)) = -\frac{D'(t)}{D(t)}.$$

In the context of reliability theory, if $S \in \mathcal{D}^*$ is a *survival function*, then r_S is the associated *failure rate*.

Corollary 1. *Let $D \in \mathcal{D}^*$. Then the following conditions are equivalent:*

(i) The discount function D exhibits (strictly) DI;

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

The non-strict part of Corollary 1 is obvious from Proposition 1 and is well known: see, for example, [27, Corollary 1]. The strict part follows straightforwardly from the fact that $\ln(D)$ is strictly convex iff r is strictly decreasing.[11]

In reliability theory, the phenomenon of decreasing failure rates is also of interest and has been much studied. We review some of this literature in Section 4 below.

2.3 Comparative DI

What does it mean to say that one decision-maker exhibits "more decreasing impatience than" another? The natural answer, proposed by Prelec ([22, Definition 2]), is embodied in the following definition:

[9]The proof of Theorem 3.3 in [14] can easily be 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 $[0, \infty)$.

[10]Recall that we use the acronym "DI" interchangeably for the noun "decreasing impatience" and the adjective "decreasingly impatient".

[11]If a function v is twice continuously 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 ([24]).

Definition 2. *We say that* $\succcurlyeq_1$ *exhibits [uniformly] 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* $(x, t+\sigma) \preccurlyeq_1 [\prec_1] (y, s+\sigma+\rho)$*.*

In fact, Prelec only proposed the "more DI" part of Definition 2. The "uniformly more DI" part is introduced here. It is a natural strengthening; additional motivation is provided by Corollary 2 (below) and the discussion following. Pratt ([21]) defines "weak" and "strong" forms of comparative risk aversion along the same lines.

Note that the sign of ρ is not restricted in Definition 2, so the comparative concept is not limited to preferences that exhibit DI.

Proposition 2 (cf. [22], Proposition 1)**.** *Let* $\succcurlyeq_1$ *and* $\succcurlyeq_2$ *be two preference orders with discounted utility representations* (u_1, D_1) *and* (u_2, D_2)*, respectively. The following conditions are equivalent:*

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

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

Proposition 2 is proved in Appendix A.2. We follow Prelec's ([22]) proof of his Proposition 1, with suitable adjustments to accommodate the uniformly more DI case. The required adjustments are not substantial but we have included a proof as it clarifies some details omitted from Prelec's original argument.

Condition (ii) is equivalent to $\ln(D_1) = h \circ \ln(D_2)$ for some strictly increasing and (strictly) convex transformation h. Since the utility functions u_1 and u_2 are irrelevant to the comparative DI properties of the preference relations, we say that the discount function D_1 exhibits (uniformly) more DI than the discount function D_2 if $\ln(D_1)$ is a strictly increasing and (strictly) convex transformation of $\ln(D_2)$. We write $D_1 \succcurlyeq_{DI} D_2$ ($D_1 \gg_{DI} D_2$) if D_1 exhibits (uniformly) more DI than D_2, and use $\sim_{DI}$ to denote the symmetric part of $\succcurlyeq_{DI}$. As this notation suggests, $\gg_{DI}$ is **not** the asymmetric part of $\succcurlyeq_{DI}$. It is possible that $D_1 \succcurlyeq_{DI} D_2$ but neither $D_1 \sim_{DI} D_2$ nor $D_1 \gg_{DI} D_2$, since there exist convex functions which are neither affine nor strictly convex. The asymmetric part of $\succcurlyeq_{DI}$ will be denoted $\succ_{DI}$.

We note two important corollaries of Proposition 2. First, for an exponential discount function, $D_2(t) = \exp(-\rho t)$ for some $\rho > 0$, we have $\ln D_2(t) = -\rho$. It follows, as Prelec ([22]) observes, that D_1 is more DI than D_2 iff D_1 is log-convex. We may likewise conclude that D_1 is *uniformly* more DI than D_2 iff D_1 is *strictly* log-convex. This fact helps to motivate our introduction of the "uniformly more DI than" relation.

Corollary 2 (cf., [22])**.** *Let* $\rho > 0$*. The discount function* D *exhibits (strictly) DI if and only if* D *exhibits (uniformly) more DI than* $\exp(-\rho t)$*.*

Second, a strictly increasing function $h : (-\infty, 0] \to (-\infty, 0]$ is affine if and only if h and h^{-1} are both convex. Since all discount functions satisfy $D(0) = 1$ we have:

Corollary 3 ([22])**.** *The discount functions D_1 and D_2 are equally DI (i.e., $D_1 \sim_{DI} D_2$) if and only if there exists some $c > 0$ such that $D_1(t) = D_2(t)^c$ for all t.*

In the literature on reliability theory, a subset of $\mathcal{D}$ (interpreted as a set of survival functions) is called a *Lehmann family* (or *proportional hazard family*) if it is of the form $\{ S^c \mid c > 0 \}$ for some $S \in \mathcal{D}$. A Lehmann family is therefore an equivalence class for the $=_{DI}$ relation (or "DI equivalence class"). The set of exponential functions in $\mathcal{D}$ is one such class.

Prelec ([22]) also constructs a local *index of decreasing impatience* for discount functions in $\mathcal{D}^*$. It measures the local degree of log-convexity and is analogous to the Arrow-Pratt index of risk aversion for "smooth" vNM utility functions. The index of DI for $D \in \mathcal{D}^*$ is denoted I_D and defined as follows:

$$I_D(t) \;=\; -\frac{r_D'(t)}{r_D(t)} \;=\; -\frac{d}{dt}\ln\left[r_D(t)\right] \;=\; -\frac{D''(t)}{D'(t)} + \frac{D'(t)}{D(t)} \qquad (2)$$

The following result confirms that this is the "right" definition.

Proposition 3 (cf., Proposition 2 in [22])**.** *Let D_1 and D_2 be discount functions in $\mathcal{D}^*$. Then $D_1 \succcurlyeq_{DI} D_2$ if and only if $I_{D_1} \geq I_{D_2}$. Moreover, $D_1 \gg_{DI} D_2$ if and only if $D_1 \succcurlyeq_{DI} D_2$ and the set*

$$\{ t \in T \mid I_{D_1}(t) = I_{D_2}(t) \} \qquad (3)$$

contains no open subset (other than the empty set).

Proof. Prelec [22, Proposition 2] proves the first claim by applying the "weak" part of Pratt's [21, Theorem 1] – specifically, the equivalence between (a) and (d). The second claim follows by applying the "strong" part of the same theorem. □

From Propositions 1 and 3 it follows that $D \in \mathcal{D}^*$ is DI if and only if $I_D \geq 0$, and is strictly DI if and only if $I_D \geq 0$ and $\{ t \mid I_D(t) = 0 \}$ contains no non-empty open set.[12] In particular, the index of DI for any exponential discount function is identically zero.

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

Example 1. *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:*

$$r_D(t) = \alpha(1 + ht)^{-1} \quad \text{and} \quad I_D(t) = h(1 + ht)^{-1}.$$

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 is [uniformly] more DI than D_2 if and only if $h_1 \geq [>] h_2$. The parameter h may therefore be used as a measure of the degree of DI of a generalized hyperbolic discount function. The parameter α has no influence on I_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.*

[12] Similarly, D is II if and only if $I_D \leq 0$ and strictly II if and only if $I_D \leq 0$ and the empty set is the only open set contained in $\{ t \mid I_D(t) = 0 \}$.

Finally, let us observe the following transitivity properties of the comparative DI relations.

Proposition 4. *The $\succcurlyeq_{DI}$ binary relation is transitive. Moreover, if $D_1 \gg_{DI} D_2$ and $D_2 \succcurlyeq_{DI} D_3$, or $D_1 \succcurlyeq_{DI} D_2$ and $D_2 \gg_{DI} D_3$, then $D_1 \gg_{DI} D_3$. In particular, $\gg_{DI}$ is transitive.*

Proof. The claims follow straightforwardly by combining Proposition 2 with Lemma 2 in Appendix A.1: the composition of two convex, strictly increasing functions is convex and strictly increasing, and is strictly convex if at least one of the functions being composed is strictly convex. □

3 Mixtures of discount functions

We now turn to the main focus of the paper: mixtures of discount (or survival) functions. We wish to use the notion of comparative DI to compare such mixtures to their components. As described in the Introduction, mixtures of discount and survival functions arise naturally in a number of contexts.

Given a set of functions $\{D_1, D_2, \ldots D_n\} \subseteq \mathcal{D}$, we define a *mixture* of them is a function

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

with $0 < \lambda_i < 1$ for all i and $\sum_{i=1}^{n} \lambda_i = 1$. Note that our definition requires that each D_i has a *strictly* positive weight. Importantly, the sets $\mathcal{D}$ and $\mathcal{D}^*$ are both closed under the formation of such mixtures.

Our starting point is the following important result, again due to Prelec ([22, Corollary 4]):

Proposition 5 (cf., [22])**.** *If D is a mixture of $\{D_1, D_2\} \subseteq \mathcal{D}$, where $D_1 \neq D_2$ and $D_1 \sim_{DI} D_2$, then $D \gg_{DI} D_2$.*

Prelec's result only asserts the weaker consequent that $D \succcurlyeq_{DI} D_2$ but the stronger conclusion follows easily from his proof.

It is obviously of interest to know whether Proposition 5 can be extended to more than two DI-equivalent functions, and also whether anything can be said about mixtures of functions that are DI-comparable but do not necessarily lie within the same DI equivalence class. To answer these questions, we make use of the following (modest) extension of Artin's result that the sum of log-convex functions is log-convex ([2, Theorem 1.8]).[13]

Lemma 1. *Let $I \subseteq \mathbb{R}_{++}$ be a non-empty convex set. If $f_i : I \to \mathbb{R}$ is log-convex for every $i \in \{1, ., ..., n\}$ (and at least one f_i is strictly log-convex), then $g = \sum_{i=1}^{n} f_i$ is (strictly) log-convex. Moreover, if g is also log-concave then there exist $k_i \in \mathbb{R}$ such that $f_i = k_i g$ for each $i \in \{1, ., ..., n\}$.*

[13] We suspect that the content of Lemma 1 is well known to many mathematicians. Certainly the arguments are standard. We include a proof for completeness, as we have not been able to find this particular result stated explicitly elsewhere.

Proof. Let $x, y \in I$ such that $x \neq y$ and let $\lambda \in (0,1)$. Using the fact that each f_i is log-convex we have:

$$g(\lambda x + (1-\lambda)y) \;=\; \sum_{i=1}^{n} f_i(\lambda x + (1-\lambda)y) \;\leq\; \sum_{i=1}^{n} f_i(x)^{\lambda} f_i(y)^{1-\lambda}. \tag{4}$$

The inequality in (4) is strict if at least one f_i is strictly log-convex. By the Weighted AM-GM Inequality [9, Theorem 7.6]:

$$\sum_{i=1}^{n} \left[\frac{f_i(x)}{g(x)}\right]^{\lambda} \left[\frac{f_i(y)}{g(y)}\right]^{1-\lambda} \;\leq\; \sum_{i=1}^{n} \left[\lambda \frac{f_i(x)}{g(x)} + (1-\lambda)\frac{f_i(y)}{g(y)}\right] \;=\; 1 \tag{5}$$

Moreover, equality holds in (5) iff

$$\frac{f_i(x)}{g(x)} \;=\; \frac{f_i(y)}{g(y)} \tag{6}$$

for all i. Combining (4) and (5) we have

$$g(\lambda x + (1-\lambda)y) \;\leq\; g(x)^{\lambda} g(y)^{1-\lambda} \tag{7}$$

which proves that g is log-convex. Since (4) holds strictly whenever some f_i is strictly log-convex, it follows that g is strictly log-convex in this case also.

Suppose g is both log-convex and log-concave. It follows that equality holds in both (4) and (5) for any $x, y \in I$ and any $\lambda \in (0,1)$. Therefore (6) holds for all i, any $x, y \in I$ and any $\lambda \in (0,1)$. In other words, f_i/g is constant on I for any i. □

Corollary 4. *If D is a mixture of $\{D_1, D_2, ..., D_n\} \subseteq \mathcal{D}$ and $D_i \succcurlyeq_{DI} D_{i+1}$ for each $i \in \{1, 2, .., n-1\}$ then $D \succcurlyeq_{DI} D_n$. Moreover, if $D_i \gg_{DI} D_{i+1}$ for some $i \in \{1, 2, .., n-1\}$ then $D \gg_{DI} D_n$.*

Proof. Let $D = \sum_{i=1}^{n} \lambda_i D_i$. Then

$$D\left(D_n^{-1}(e^z)\right) \;=\; \sum_{i=1}^{n-1} \lambda_i D_i\left(D_n^{-1}(e^z)\right) + \lambda_n e^z.$$

Using Proposition 4, we know that $D_i \succcurlyeq_{DI} D_n$ for all $i < n$, so $D_i\left(D_n^{-1}(e^z)\right)$ is log-convex for all $i < n$. Moreover, if $D_i \gg_{DI} D_{i+1}$ then $D_i \gg_{DI} D_n$ (Proposition 4) so $D_i\left(D_n^{-1}(e^z)\right)$ is *strictly* log-convex. It follows that $\lambda_i D_i\left(D_n^{-1}(e^z)\right)$ is log-convex for all $i < n$ and strictly log-convex if $D_i \gg_{DI} D_{i+1}$. Since $\lambda_n e^z$ is obviously log-convex, the result now follows by Lemma 1 and Proposition 2. □

Corollary 5. *If D is a mixture of $\{D_1, D_2, ..., D_n\} \subseteq \mathcal{D}$ and each D_i exhibits DI then D exhibits DI. Moreover, if at least one D_i exhibits strictly DI then D exhibits strictly DI.*

Proof. Let $D_{n+1}(t) = e^{-t}$ for all $t \in T$ and let $D = \sum_{i=1}^{n} \lambda_i D_i$. Then

$$D\left(D_{n+1}^{-1}(e^z)\right) \;=\; \sum_{i=1}^{n} \lambda_i D_i\left(D_{n+1}^{-1}(e^z)\right).$$

Each $D_i\left(D_{n+1}^{-1}(e^z)\right)$ is log-convex by Proposition 2 and Corollary 2, so each $\lambda_i D_i\left(D_{n+1}^{-1}(e^z)\right)$ is log-convex. Therefore, $D\left(D_{n+1}^{-1}(e^z)\right)$ is log-convex by Lemma 1, so D exhibits DI (Proposition 2 and Corollary 2). Moreover, if some D_i exhibits strictly DI then $D\left(D_{n+1}^{-1}(e^z)\right)$ is strictly log-convex so D exhibits strictly DI. □

3.1 Mixing equally-DI functions

We can now extend Proposition 5 to mixtures of any finite number of functions.

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

Proof. Proposition 5 establishes the result for the case $n = 2$. Suppose that $k \geq 2$ and the result is true for all $n \leq k$. Let D be a mixture of $D_1, D_2, \ldots, D_{k+1}$ with $D_j \in \mathcal{D}$ for each j, $D_1 \neq D_j$ for some $j \leq k+1$ and $D_1 \sim_{DI} D_2 \sim_{DI} \ldots \sim_{DI} D_{k+1}$. We therefore have:

$$D \;=\; \sum_{i=1}^{k+1} \lambda_i D_i \;=\; (1-\lambda_{k+1})\left[\sum_{i=1}^{k}\left(\frac{\lambda_i}{1-\lambda_{k+1}}\right) D_i\right] + \lambda_{k+1} D_{k+1} \tag{8}$$

If $D_1 \neq D_j$ for some $j \leq k$ then

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

by the inductive hypothesis, and hence

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

by Proposition 4. On the other hand, if $D_1 = D_j$ for all $j \leq k$, then

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

and $D_k \neq D_{k+1}$, so

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

by Proposition 5. In either case, we deduce $D \gg_{DI} D_{k+1}$ using (8) and Corollary 4. □

From Corollary 3 we know that any two exponential discount functions are equally DI. It follows from Theorem 1 that mixtures of non-identical exponential functions will be uniformly more DI than each component function. This is a continuous-time version of Jackson and Yariv's result ([15, Proposition 1]) and is illustrated by the following example.

Example 2. *Let* $D_i(t) = \exp(-\rho_i t)$, *with* $\rho_1 = 0.01, \rho_2 = 0.02, \rho_3 = 0.03$. *Consider their mixture* $D = (1/3)(D_1 + D_2 + D_3)$. *Then* $I_{D_i}(t) = 0$ *for all* t *whereas the Index of DI for the mixture is:*

$$I_D = \frac{e^{-(\rho_1+\rho_2)t}(\rho_1 - \rho_2)^2 + e^{-(\rho_2+\rho_3)t}(\rho_2 - \rho_3)^2 + e^{-(\rho_1+\rho_3)t}(\rho_1 - \rho_3)^2}{\left(\rho_1 e^{-\rho_1 t} + \rho_2 e^{-\rho_2 t} + \rho_3 e^{-\rho_3 t}\right)\left(e^{-\rho_1 t} + e^{-\rho_2 t} + e^{-\rho_3 t}\right)}.$$

Clearly, this is strictly greater than zero everywhere, as also shown in Figure 1.

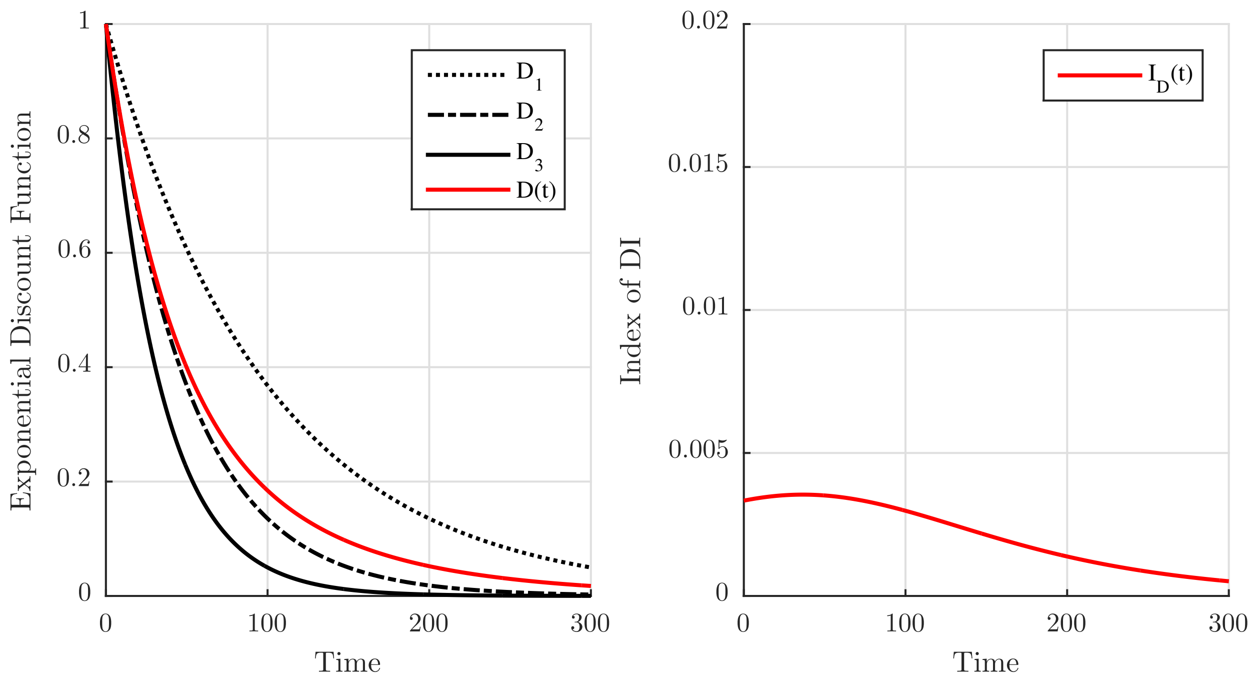

Figure 1: Index of DI for the Mixture of Exponential Discount Functions

However, there are other interesting DI equivalence classes besides the one containing the exponential functions. Recalling Example 1, Theorem 1 implies that if we mix generalised hyperbolic discount functions with the same h parameter but different α values, then the resulting mixture will be uniformly more DI than each component.

Example 3. *Let* $D_i(t) = (1 + ht)^{-\alpha_i/h}$, *with* $\alpha_1 = 1, \alpha_2 = 2, \alpha_3 = 3, h = 5$. *Consider their mixture* $D = (1/3)(D_1 + D_2 + D_3)$. *Recall from Example 1 that* $I_{D_i}(t) = h_i(1 + h_i t)^{-1} = 5(1 + 5t)^{-1}$. *From Figure 2 it can be seen that* I_D *lies strictly above* I_{D_i}.

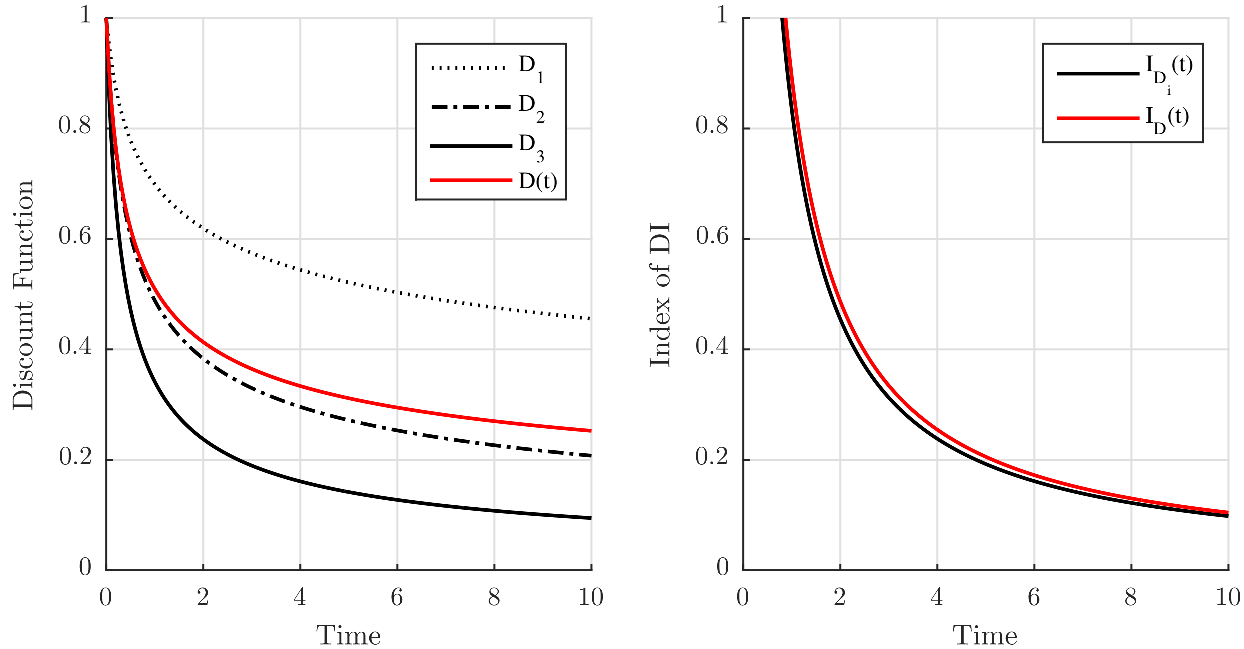

Figure 2: Index of DI for the Mixture of Generalized Hyperbolic Discount Functions

3.2 Mixing weakly DI-ordered functions

If we mix $D_2 \in \mathcal{D}$ with some $D_1 \in \mathcal{D}$ for which $D_1 \sim_{DI} D_2$, we obtain a mixture that is uniformly more DI than D_2. It is natural to expect, therefore, that if we mix D_2 with some $\hat{D}_1$ that is *more* DI than D_1, the resulting mixture will also be uniformly more DI than D_2. This expectation is predicated on an implicit assumption that the "more DI than" relation will satisfy the following *independence* property for any $\lambda \in (0,1)$:

$$\hat{D}_1 \succcurlyeq_{DI} D_1 \quad \Rightarrow \quad \lambda \hat{D}_1 + (1-\lambda) D_2 \succcurlyeq_{DI} \lambda D_1 + (1-\lambda) D_2$$

In fact, it does not.

Proposition 6. *Let $D_1, D_2 \in \mathcal{D}$ with $D_1 \neq D_2$ and $D_1 \succcurlyeq_{DI} D_2$. Then there exists $\hat{D}_1 \in \mathcal{D}$ such that $D_1 \succcurlyeq_{DI} \hat{D}_1 \succcurlyeq_{DI} D_2$ and*

$$\lambda \hat{D}_1 + (1-\lambda) D_2 \gg_{DI} D_2$$

for any $\lambda \in (0,1)$. However, there exist $D_1, D_2 \in \mathcal{D}$ with $D_1 \neq D_2$ and $D_1 \succcurlyeq_{DI} D_2$, and $\lambda \in (0,1)$, for which it is **not** *the case that*

$$\lambda D_1 + (1-\lambda) D_2 \gg_{DI} D_2.$$

Proof. Consider the first claim. Let $\hat{D}_1 = D_2^c$ for some $c > 0$ with $c \neq 1$. Then $\hat{D}_1 \neq D_2$ and $\hat{D}_1 \sim_{DI} D_2$ (Corollary 3) so any mixture of $\hat{D}_1$ and D_2 will be uniformly more DI than D_2 (Proposition 5). It remains to show that $D_1 \succcurlyeq_{DI} \hat{D}_1$. We know that there exists a convex, strictly increasing function, $h : (-\infty, 0] \to (-\infty, 0]$ such that $\ln D_1 = h \circ \ln D_2$ (Proposition 2). Therefore

$$\ln D_1 = h \circ \left[\frac{1}{c} \ln \hat{D}_1 \right] = \hat{h} \circ \ln \hat{D}_1$$

where $\hat{h} = h \circ g$ and $g(z) = c^{-1}z$ for all $z \in (-\infty, 0]$. Since $\hat{h}$ is convex (Lemma 2 in Appendix A.1) and strictly increasing, we deduce that $D_1 \succcurlyeq_{DI} \hat{D}_1$ (Proposition 2).

To establish the second claim, consider the following example. Let $\rho_1 > \rho_2 > 0$ and let $\bar{t} > 0$. Define $D_2(t) = \exp(-\rho_2 t)$ and

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

where $\overline{d} = (\rho_1 - \rho_2)\bar{t} > 0$. It is easily checked that $D_1 \in \mathcal{D}$. (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

$$\ln D_1\left(D_2^{-1}(e^z)\right) = \begin{cases} (\rho_1/\rho_2)\, z & \text{if } z > -\rho_2 \bar{t} \\ z - \overline{d} & \text{if } z \leq -\rho_2 \bar{t} \end{cases} \tag{9}$$

Since

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

(9) is continuous. From this fact and $(\rho_1/\rho_2) > 1$ we deduce that (9) is convex, but neither affine nor strictly convex. It follows that $D_1 \succ_{DI} D_2$.

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

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

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

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

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 $D \gg_{DI} D_2$. $\square$

This leaves open the obvious question: When is the mixture of weakly DI-ordered (and non-identical) discount functions uniformly more DI than the least-DI of the functions being mixed? The following result provides the answer.

Theorem 2. *Let $n \geq 2$ and let $D_1, D_2, \ldots, D_n$ be functions in $\mathcal{D}$ such that*

$$D_1 \succcurlyeq_{DI} D_2 \succcurlyeq_{DI} \ldots \succcurlyeq_{DI} D_n.$$

Precisely one of the following obtains:

(i) $D \gg_{DI} D_n$ for any mixture D of $\{D_1, D_2, \ldots, D_n\}$.

(ii) There exists a non-empty interval $J = (t_1, t_2) \subseteq T$ and constants $c_1, c_2, \ldots, c_{n-1}$ such that $D_i(t) = c_i D_n(t)$ for each $i \in \{1, 2, \ldots, n-1\}$ and every $t \in J$.

Proof. By Corollary 4, $D \succcurlyeq_{DI} D_n$ for any mixture D of $\{D_1, D_2, \ldots, D_n\}$. Suppose there is some such mixture $D = \sum_{i=1}^{n} \lambda_i D_i$ for which D is **not** uniformly more DI than D_n. Let $g = D\left(D_n^{-1}\left(e^z\right)\right)$. It follows that g is log-convex but not strictly log-convex. Since g is also strictly increasing and continuous, it follows that there exists a non-empty open interval $I \subseteq (-\infty, 0]$ on which $\ln g$ is affine. That is, there exist constants $\alpha > 0$ and β such that $\ln g(z) = \alpha z + \beta$ for all $z \in I$. Next, observe that g is the sum of log-convex functions:

$$g(z) \;=\; \sum_{i<n} \lambda_i D_i \left(D_n^{-1}\left(e^z\right)\right) + \lambda_n e^z$$

The function $\lambda_n e^z$ is obviously log-convex; the fact that $D_i \succcurlyeq_{DI} D_n$ (Proposition 4) implies $D_i\left(D_n^{-1}\left(e^z\right)\right)$ is log-convex (Proposition 2) and hence $\lambda_i D_i\left(D_n^{-1}\left(e^z\right)\right)$ is log-convex. Let $f_i(z) = \lambda_i D_i\left(D_n^{-1}\left(e^z\right)\right)$ if $i < n$ and let $f_n(z) = \lambda_n e^z$. In particular, from Proposition 2,

$$\ln\left(D_i(t)\right) = \ln f_i\left(\ln\left(D_n(t)\right)\right) - \ln\left(\lambda_i\right) \tag{10}$$

for any $i < n$ and any $t \in J$. Since g is affine on I it follows by Lemma 1 that there exist constants $k_1, k_2, \ldots, k_n$ such that

$$\ln f_i(z) \;=\; \ln\left(k_i\right) + \ln g(z) \;=\; \alpha z + \left[\beta + \ln\left(k_i\right)\right]$$

for all i and all $z \in I$. The case $i = n$ gives

$$\ln\left(\lambda_n\right) + z \;=\; \alpha z + \left[\beta + \ln\left(k_n\right)\right]$$

for all $z \in I$. It follows that $\alpha = 1$ and hence, for any $i < n$,

$$\ln f_i(z) \;=\; z + d_i$$

for any $z \in I$, where $d_i = \beta + \ln\left(k_i\right)$. From (10) we now deduce that

$$\ln D_i(t) \;=\; \ln D_n(t) + \hat{d}_i$$

for any $i < n$ and any $t \in J$, where $\hat{d}_i = d_i - \ln\left(\lambda_i\right)$. Letting $c_i = \exp\left(\hat{d}_i\right)$ does the needful.

Conversely, suppose $D_i(t) = c_i D_n(t)$ for each $i \in \{1, 2, \ldots, n-1\}$ and every t in some non-empty open interval J. Hence, $\ln D_n(J)$ is also a non-empty open interval. Let $D = \sum_{i=1}^n \lambda_i D_i$ be a mixture of $\{D_1, D_2, \ldots, D_n\}$. Then, for any $z \in \ln D_n(J)$ we have:

$$\begin{aligned} D\left(D_n^{-1}\left(e^z\right)\right) &= \sum_{i<n} \lambda_i D_i \left(D_n^{-1}\left(e^z\right)\right) + \lambda_n e^z \\ &= \sum_{i<n} \lambda_i c_i D_n \left(D_n^{-1}\left(e^z\right)\right) + \lambda_n e^z \\ &= \left[\lambda_n + \sum_{i<n} \lambda_i c_i\right] e^z \end{aligned}$$

In other words, $\ln D\left(D_n^{-1}\left(e^z\right)\right)$ is affine on the interval $\ln D_n(J)$, and therefore $\ln D$ is not a strictly convex transformation of $\ln D_n$. It follows (Proposition 2) that D is not uniformly more DI than D_n. □

If the elements of $\{D_1, D_2, \ldots, D_n\}$ are in the same DI equivalence class, Corollary 3 implies that there exist constants $\alpha_1, \alpha_2, \ldots, \alpha_{n-1}$ such that $D_i = D_n^{\alpha_i}$ for each $i < n$. Unless $\alpha_i = 1$ for each $i < n$, this will exclude case (ii) of Theorem 2. In other words, Theorem 1 is consistent with Theorem 2.

Example 4. *Let $D_i(t) = (1 + h_i t)^{-1}$, with $h_1 = 0.03, h_2 = 0.02, h = 0.01$. These are* proportional hyperbolic *functions. Recalling Example 1, we see that $D_1 \succcurlyeq_{DI} D_2 \succcurlyeq_{DI} D_3$. Consider their mixture $D = (1/3)(D_1 + D_2 + D_3)$. It is obvious that*

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

is strictly increasing (decreasing) in t when $h_j > h_i$ ($h_j < h_i$). From Figure 3 it can be seen that I_D lies strictly above I_{D_3}.

Theorem 1 also suggests the possibility that mixtures of (strictly) DI-ordered discount functions may be uniformly more DI than the *most* DI of the functions being mixed. Examples are not difficult to find, though a general characterisation of this phenomenon remains elusive.

Example 5. *The following example is based on [13]. Let $D_1(t) = \exp(-t)$ and let $D_2(t) = \exp\left(-\rho\left(e^t - 1\right)\right)$ with $\rho > 1$. Hence, D_1 is exponential and D_2 is a truncated extreme value discount function. By straightforward calculation, $I_{D_1}(t) = 0$ and $I_{D_2}(t) = -1$. It follows (Proposition 3) that $D_1 \gg_{DI} D_2$. In particular, D_2 exhibits strictly* increasing *impatience (or a strictly increasing failure rate in the language of reliability theory). Let $D = \lambda_1 D_1 + \lambda_2 D_2$ be a mixture of D_1 and D_2.*

Gurland and Sethuraman ([13]) consider the case $\rho = 20$ and show by numerical simulation that r_D may be strictly decreasing (for sufficiently high values of λ_1), which implies $D \gg_{DI} D_1$ (Corollary 2). In fact, it is not hard to show that r_D is strictly decreasing iff

$$\lambda_1 \geq \frac{\rho}{(\rho - 1)^2}$$

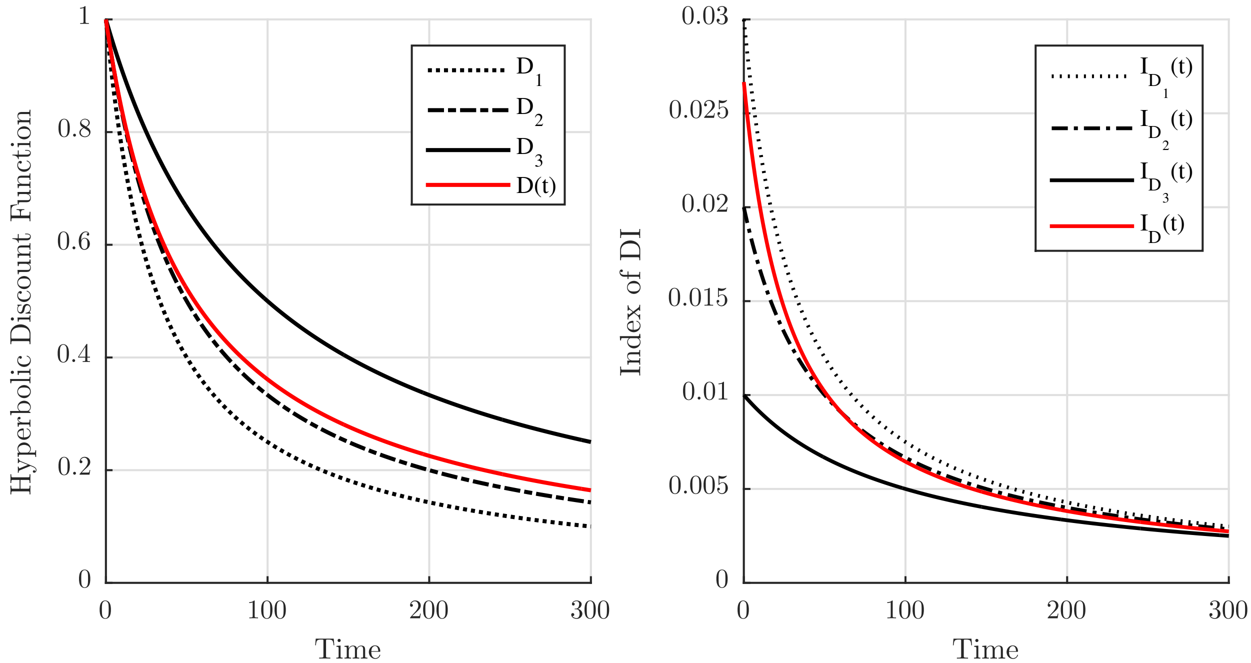

Figure 3: Index of DI for the Mixture of Generalized Hyperbolic Discount Functions

(The value of the RHS is contained in the interval $(0, 1)$ *provided* ρ *is high enough.) We relegate the details to Appendix A.3.*

The key point to observe from this example is that the necessary and sufficient condition for r_D *to be strictly decreasing is a* joint *restriction on* ρ *and the mixing parameters. This precludes the possibility of specifying conditions on an arbitrary set of weakly DI-ordered functions that are necessary and sufficient for* **any** *mixture to uniformly dominate the most-DI component. At best, we can establish sufficient conditions.*

3.3 A general result for mixing in $\mathcal{D}^*$

Theorem 2 requires that the functions being mixed are weakly DI-ordered, which places significant restrictions on its applicability. However, we can develop a "local" analogue of this result which applies to mixtures of *any* finite set of "smooth" discount functions. Suppose D is a mixture of $\{D_1, D_2, \ldots, D_n\} \subseteq \mathcal{D}^*$. The indices $\{I_{D_1}(t), I_{D_2}(t), \ldots, I_{D_n}(t)\}$ can obviously be ordered at any t. It is therefore natural to seek necessary and sufficient conditions for $I_D(t)$ to strictly exceed

$$\min\{I_{D_1}(t), \ldots, I_{D_n}(t)\}.$$

Theorem 3. *Let* $\{D_1, D_2, \ldots, D_n\} \subseteq \mathcal{D}^*$ *and let*

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

be a mixture of $D_1, D_2, \ldots, D_n$. *Then* $I_D \geq \min_i\{I_{D_i}\}$ *on* $[0, \infty)$ *and*

$$I_D(t) = \min_i\{I_{D_i}(t)\}$$

iff $r_{D_1}(t) = r_{D_2}(t) = \cdots = r_{D_n}(t)$ *and* $r'_{D_1}(t) = r'_{D_2}(t) = \cdots = r'_{D_n}(t)$.

Proof. Let $r_i = r_{D_i}$ and $I_i = I_{D_i}$ for each i. Recall that $D' = -rD$ and hence

$$D'' = Dr_D^2 - Dr'_D = Dr_D(r_D + I_D).$$

Recall also that

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

Therefore,

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

This expression can be rearranged as follows:

$$I_D \;=\; \frac{\sum_{i=1}^n \lambda_i D_i r_i I_i}{\sum_{i=1}^n \lambda_i D_i r_i} + \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i} \;=\; \sum_{i=1}^n \alpha_i I_i + Q,$$

where

$$Q \;=\; \frac{\sum_{i=1}^n \lambda_i D_i r_i^2}{\sum_{i=1}^n \lambda_i D_i r_i} - \frac{\sum_{i=1}^n \lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i}$$

and

$$\alpha_i \;=\; \frac{\lambda_i D_i r_i}{\sum_{i=1}^n \lambda_i D_i r_i}$$

Note that $\sum_{i=1}^n \alpha_i(t) = 1$ and $\alpha_i(t) > 0$ for all t.

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. 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:

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

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

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

It follows that $Q(t) \geq 0$ for all t, with equality iff $r_1(t) = r_2(t) = \cdots = r_n(t)$. Since

$$\min_i \{I_i\} \leq \min_i \{I_i\} + Q \leq \sum_{i=1}^{n} \alpha_i I_i + Q = I_D.$$

we have $I_D \geq \min_i \{I_i\}$ on $[0, \infty)$ and $I_D(t) > \min_i \{I_i(t)\}$ if $r_j(t) \neq r_k(t)$ for some $j \neq k$. Conversely, if $r_1(t) = r_2(t) = \cdots = r_n(t)$ then $Q(t) = 0$ and

$$I_D(t) = \sum_{i=1}^{n} \alpha_i(t) I_i(t).$$

In this case, $I_D(t) = \min_i I_i(t)$ if and only if $I_1(t) = I_2(t) = \cdots = I_n(t)$, which holds if and only if $r'_1(t) = r'_2(t) = \cdots = r'_n(t)$. □

Corollary 6. *Let $\{D_1, D_2, \ldots, D_n\} \in \mathcal{D}^*$. Assume that $D_1, D_2, \ldots, D_n$ all exhibit DI. Let*

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

be a mixture of $D_1, D_2, \ldots, D_n$. Then $I_D(t) \geq 0$ for all t, with equality iff

$$r_{D_1}(t) = r_{D_2}(t) = \ldots = r_{D_n}(t) \quad \text{and} \quad r'_{D_1}(t) = r'_{D_2}(t) = \ldots = r'_{D_n}(t) = 0.$$

Hence, D exhibits DI, and strictly DI unless the set

$$\left\{ t \in T \mid r_{D_1}(t) = r_{D_2}(t) = \ldots = r_{D_n}(t) \text{ and } r'_{D_1}(t) = r'_{D_2}(t) = \ldots = r'_{D_n}(t) = 0 \right\}$$

contains a non-empty open subset.

Proof. Recalling the notation from the previous proof, we have, by assumption, $I_i \geq 0$ for each i. It follows from Theorem 3 that $I_D \geq 0$ and hence that D exhibits DI. It exhibits strictly DI unless the set

$$\{ t \in T \mid I_D(t) = 0 \} = \left\{ t \in T \;\middle|\; \sum_{i=1}^{n} \alpha_i(t) I_i(t) + Q(t) = 0 \right\}$$

contains a non-empty open subset. Since $I_i(t) \geq 0$ and $\alpha_i(t) > 0$ for all i and all t, this is equivalent to the requirement that the set

$$\{ t \in T \mid I_i(t) = Q(t) = 0 \text{ for all } i \}$$

contain a non-empty open subset. Since $I_i(t) = 0$ iff $r'_i(t) = 0$ and $Q(t) = 0$ iff $r_i(t) = r_j(t)$ for all $i, j \in \{1, 2, \ldots, n\}$, the result follows. □

The following example provides an illustration of Theorem 3.

Example 6. *Let $n = 2$, let $D_1(t) = (1 + ht)^{-2}$ be an hyperbolic discount function and let $D_2(t) = \exp(-\alpha t^{1/2})$ a Weibull discount function [13]. Then $I_{D_1}(t) = h(1 + ht)^{-1} > 0$ for all t (Example 1) and, by straightforward calculation, $I_{D_2}(t) = (2t)^{-1} > 0$ for all t. Therefore, both D_1 and D_2 exhibit strict DI. Observe that*

$$I_{D_1}(t) - I_{D_2}(t) \;=\; \frac{ht - 1}{2t\,(1 + ht)},$$

so

$$I_{D_1}(t) \gtreqless I_{D_2}(t) \;\; as \;\; t \gtreqless h^{-1}.$$

It follows that D_1 and D_2 both are from incomparable DI classes. Let $D = \frac{1}{2}(D_1 + D_2)$. The behaviour of I_D with parameters $h = 0.1$ and $\alpha = 0.12$ is illustrated in Figure 4. It can be clearly seen that neither $D \gg_{DI} D_1$ nor $D \gg_{DI} D_2$. However, $I_D(t) > \min\{I_{D_1}(t), I_{D_2}(t)\}$ for all t. Indeed, $I_D(t) > \max\{I_{D_1}(t), I_{D_2}(t)\}$ in a large neighbourhood of $t = 10$ (the time at which I_{D_1} intersects I_{D_2}).

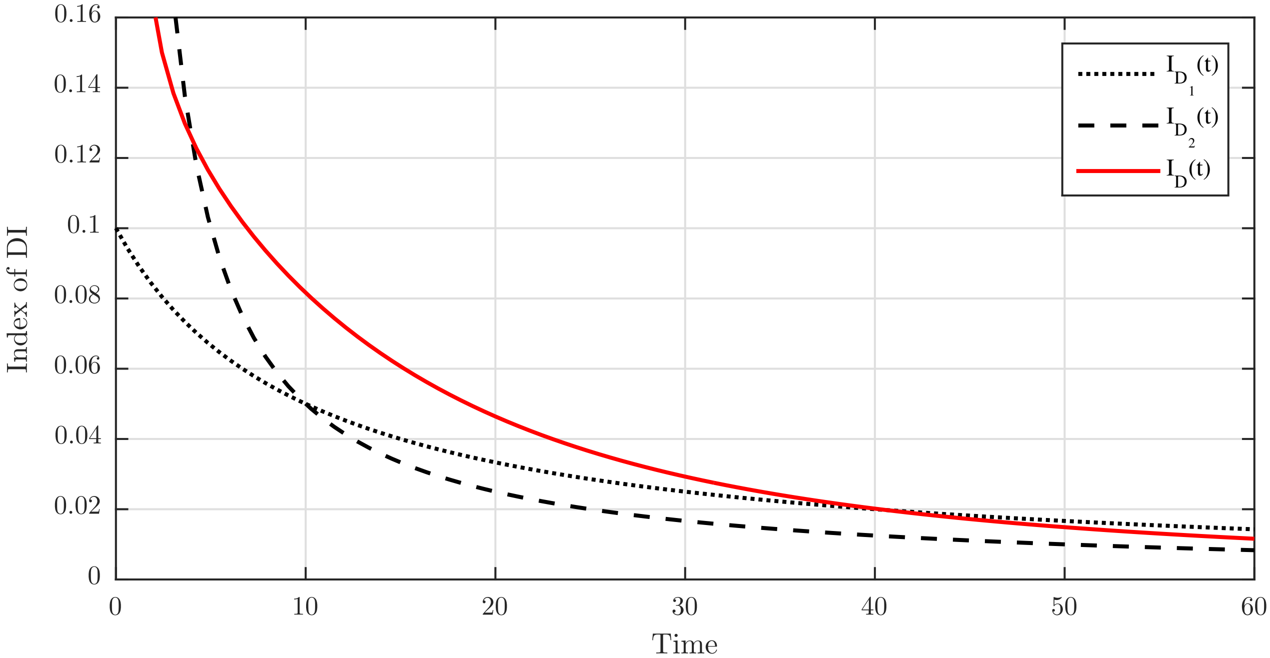

Figure 4: Index of DI for the Mixture of D_1 and D_2

4 Related results

Several results in the literature are special cases of our main theorems.

Consider Theorem 1. This generalises Corollary 4 in Prelec [22] (see Proposition 5 above). Other special cases have appeared elsewhere in the decision theory and reliability theory literatures. These results focus on mixtures of exponential functions.

Gurland and Sethuraman ([13, Example 2]) observe that the mixture of two distinct exponential survival functions exhibits a strictly decreasing failure rate. Jackson and Yariv ([15, Proposition 1]) and Pennesi ([20, Theorem 1]) are analogous results for exponential discount functions, though in discrete-time settings. While Pennesi's result establishes that the mixture satisfies the obvious discrete-time analogue of strictly decreasing impatience (called *diminishing impatience* in his paper), Jackson and Yariv establish instead that the mixture satisfies a property they call *present bias*.[14] The appropriate continuous-time analogue of this concept is not entirely unambiguous, but the following is an obvious contender:

Definition 3. *The preference order $\succcurlyeq$ exhibits present bias if*

(i) $(y, s) \succcurlyeq (x, t)$ *implies* $(y, s + \sigma) \succcurlyeq (x, t + \sigma)$ *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* $(y^*, s + \sigma) \succ (x^*, t + \sigma)$ *and* $(x^*, t) \succ (y^*, s)$.

It can be shown (see Corollary 6 in [1]) that present bias – in the sense of Definition 3 – is equivalent to strictly decreasing impatience for preferences with a DU representation. For their discrete-time environment, Jackson and Yariv's ([15]) Proposition 1 establishes that DU preferences whose associated discount function is a mixture of heterogeneous exponential functions must exhibit present bias. Our Theorem 1 therefore implies a continuous-time version of their result, and also of Pennesi's.

A somewhat weaker result is presented by Azfar ([3]). He considers an exponential discounter whose rate of time preference is the sum of a "pure" time preference rate plus the constant hazard rate from an exponential survival function describing the decision-maker's random lifespan. Suppose the decision-maker is uncertain about the true hazard rate of the survival function and suppose further that time is indexed contonuously. If r is the time preference rate for the expected discount function, Azfar proves that the "apparent discount rate"

$$\rho_a(t) = \frac{1}{t} \int_0^t r(s) ds$$

is strictly declining ([3, Proposition 1]). Gollier and Weitzman ([12]) make a similar observation. This fact is, of course, implied by the result that r itself is strictly declining.

Corollaries 5 and 6 also have antecedents in the literature. Corollary 5 shows that the mixing operation preserves the property of decreasing impatience. For mixtures of smooth functions, Corollary 6 sharpens this result by identifying a necessary and sufficient condition for the mixture to exhibit strictly DI. Proschan [23, Theorem 2] gives a direct proof

[14] Their concept differs from the eponymous notion in [20], and also from Jackson and Yariv's own use of the term in [16].

of Corollary 5 for the case of survival functions.[15] Theorem 5 in [21] is a direct analogue of our Corollary 6. Pratt considers mixtures of von Neumann-Morgenstern (vNM) utility functions that are twice continuously differentiable with positive first derivatives. A vNM utility function u exhibits "(strictly) decreasing risk aversion" if the Arrow-Pratt index of risk aversion for u,

$$\mathfrak{R}_u(x) = -\frac{d}{dx}\ln u'(x),$$

is (strictly) decreasing. Pratt proves that mixing preserves decreasing risk aversion, and that the mixture will exhibit strictly decreasing risk aversion except on sub-intervals where each vNM utility function has equal and constant indices of risk aversion.

5 Discussion

This paper has generalised and extended the important work of Prelec ([22]), with the assistance of Artin ([2]). While various special cases are known, the general results presented here seem to have escaped the attention of researchers in decision theory and reliability theory. We hope that the paper renders useful service in this respect.

In the light of Theorem 1, Proposition 6 reveals a somewhat surprising anomaly in the behaviour of mixtures of DI-ordered discount functions: the mixture need not be uniformly more DI than the least DI of the functions being mixed. The scope of this anomalous behaviour is precisely determined in Theorem 2.

Theorem 3, which is heavily influenced by the work of Pratt ([21]), provides a local analogue of our global results, for mixtures of smooth discount functions.

There remain many interesting open questions, particularly concerning conditions under which mixtures of non-DI-equivalent functions are more DI than *all* of the functions being mixed. As Example 5 makes clear, it is fruitless to seek necessary and sufficient conditions at the level of generality maintained in the present paper. We must either content ourselves with sufficient conditions, or seek more restricted domains on which necessary and sufficient conditions may be found. The reliability theory literature has partially explored both avenues ([6]), with a particular focus on mixtures of survival functions that exhibit increasing failure rates, and on conditions under which such mixtures exhibit a decreasing failure rate or specific variation in monotonicity (such as the "bathtub" shape). Conditions under which mixtures of DI functions are more DI than the most decreasingly impatient component are, to the best of our knowledge, yet to be explored.

In the decision theory context, the results presented here raise other issues. Our analysis is "static" in the sense of studying preferences elicited at a fixed point in time ("the present"). When such preferences exhibit non-constant impatience (non-stationarity), questions of dynamic inconsistency naturally arise but cannot be settled within the confines of the static model. However, as pointed out by Azfar ([3]), Pennesi ([20]), Millner and Heal ([19]) and others, decreasing impatience induced by the mixing of exponential

[15] Barlow, Marshall and Proschan ([4, Theorem 3.4]) give an even more general version of this result, which is not restricted to finite mixtures.

discount functions need not entail dynamically inconsistent choice. Provided the mixing weights are adjusted suitably over time, optimal plans can be faithfully implemented without violating the preferences of future selves.

What if the preferences being mixed are not exponential, and therefore already exhibit non-constant impatience? In this case, dynamic consistency questions arise before any mixing takes place. This opens several avenues of further enquiry. First, when can mixing induce constant impatience? In other words, can we identify a set of discount functions whose convex hull is contained within (or at least, has non-empty intersection with) the exponential class? Second, absent a stationary mixture, can the mixing weights be adjusted over time to permit dynamically consistent decision-making? Can these adjustments be justified other than by direct appeal to a dynamic consistency axiom? Finally, if mixing reflects aggregation by a Social Planner, how can the Planner determine socially optimal courses of action based on this utilitarian aggregation of non-stationary preferences? How should the preferences of the future selves of current individuals be factored into social decision-making?

Acknowledgments

We thank Matthew Jackson, John Hillas, Simon Grant and several seminar audiences for comments and suggestions. Slinko was supported by the Royal Society of New Zealand Marsden Fund 3706352. Anchugina gratefully acknowledges financial support from the University of Auckland and the Centre for Social Data Analytics (CSDA) at Auckland University of Technology, as well as the hospitality of Victoria University of Wellington during the writing of this paper.

A Appendix

A.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 [26].[16]

[16] An interval is "non-trivial" if it is neither empty nor a singleton.

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 2. *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 [8]. 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:

$$f(\lambda x + (1 - \lambda)y) \leq f(x)^{\lambda} f(y)^{1-\lambda}. \tag{11}$$

The function f is strictly log-convex if inequality (11) is strict when $x \neq y$ and $\lambda \in (0, 1)$.

One of the important definitions which will be frequently used throughout the paper 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 $f_1 = f \circ f_2$.

Lemma 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 [21]. □

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.

A.2 Proof of Proposition 2

We need to prove the following lemma first:

Lemma 4. *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

$$h_2(s) - h_2(t) = h_2(s + \sigma + \rho) - h_2(t + \sigma). \tag{12}$$

We need to show that

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

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 (12) 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:

$$h_1(s) - h_1(t) \leq [<]\, h_1(s + \sigma + \rho) - h_1(t + \sigma), \tag{13}$$

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

To show the sufficiency, suppose that (12) implies (13) 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 (12) 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 2.

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,

$$h_2(s) - h_2(t) = h_2(s+\rho+\sigma) - h_2(t+\sigma) \tag{14}$$

and

$$h_1(s) - h_1(t) > [\geq]\, h_1(s+\rho+\sigma) - h_1(t+\sigma). \tag{15}$$

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 4 this conclusion contradicts equation (14) and inequality (15). 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 4, h_1 is a (strictly) convex transformation of h_2. □

A.3 Example 5: Omitted details

Gurland and Sethuraman ([13, Corollary 2]) establish that $r'_D(t) \leq [<] 0$ iff

$$\rho e^t \leq [<] \left[\frac{Q_1(t)}{Q_1(t) + Q_2(t)}\right] \left(\rho e^t - 1\right)^2$$

where $Q_i = \lambda_i D_i$. Let

$$f(t) = \left[\frac{Q_1(t)}{Q_1(t) + Q_2(t)}\right] \left(\rho e^t - 1\right)^2 - \rho e^t$$

A *necessary* condition for r_D to be strictly decreasing is $f(0) \geq 0$, which is equivalent to

$$\lambda_1 \geq \frac{\rho}{(\rho - 1)^2} \tag{16}$$

We will show that (16) also suffices.

Using the facts that $Q_1'(t) = -Q_1(t)$ and $Q_2'(t) = -\rho e^t Q_2(t)$, straightforward calculations give:

$$f'(t) \;=\; f(t) \;+\; \left[\left(\rho e^t\right)^2 - 1\right]\left[\frac{Q_1(t)}{Q_1(t) + Q_2(t)}\right] \;+\; \left(\rho e^t - 1\right)^3 \left[\frac{Q_1(t)Q_2(t)}{\left(Q_1(t) + Q_2(t)\right)^2}\right]$$

Since

$$\left[\left(\rho e^t\right)^2 - 1\right]\left[\frac{Q_1(t)}{Q_1(t) + Q_2(t)}\right] \;+\; \left(\rho e^t - 1\right)^3 \left[\frac{Q_1(t)Q_2(t)}{\left(Q_1(t) + Q_2(t)\right)^2}\right] \;>\; 0$$

for all $t \geq 0$, we see that $f'(t) > 0$ for all $t > 0$ if $f(0) \geq 0$, which is equivalent to (16).

References

[1] N. Anchugina, M.J. Ryan, and A. Slinko. Aggregating preferences with decreasing impatience. *arXiv preprint*, arxiv: 1604.01819, 2016.

[2] E. Artin. *The gamma function.* Holt, Rinehart and Winston, New York, 1964.

[3] O. Azfar. Rationalizing hyperbolic discounting. *Journal of Economic Behavior & Organization*, 38(2):245–252, 1999.

[4] R.E. Barlow, A.W. Marshall, and F. Proschan. Properties of probability distributions with monotone hazard rate. *Annals of Mathematical Statistics*, 34(2):375–389, 1963.

[5] H. Block and H. Joe. Tail behavior of the failure rate functions of mixtures. *Lifetime Data Analysis*, 3(3):269–288, 1997.

[6] H. W. Block. The failure rates of mixtures. *Advances in Distribution Theory, Order Statistics, and Inference*, pages 267–277, 2006.

[7] H.W. Block, Y. Li, and T.H. Savits. Initial and final behaviour of failure rate functions for mixtures and systems. *Journal of Applied probability*, 40(3):721–740, 2003.

[8] S. Boyd and L. Vandenberghe. *Convex optimization.* Cambridge University Press, Cambridge, 2004.

[9] Z. Cvetkovski. *Inequalities: Theorems, Techniques and Selected Problems*. Springer-Verlag, Berlin Heidelberg, 2012.

[10] P. C. Fishburn and A. Rubinstein. Time preference. *International Economic Review*, 23(3):677–694, 1982.

[11] 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.

[12] C. Gollier and M. L. Weitzman. How should the distant future be discounted when discount rates are uncertain? *Economics Letters*, 107:350–353, 2010.

[13] 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.

[14] C. M. Harvey. Proportional discounting of future costs and benefits. *Mathematics of Operations Research*, 20(2):381–399, 1995.

[15] M. O. Jackson and L. Yariv. Present bias and collective dynamic choice in the lab. *the American Economic Review*, 104(12):4184–4204, 2014.

[16] M. O. Jackson and L. Yariv. Collective dynamic choice: the necessity of time inconsistency. *American Economic Journal: Microeconomics*, 7(4):150–178, 2015.

[17] L. Lee. Distributions ordered by their hazard functions. Technical report, VPI, 1981.

[18] S.M. McClure, K.M. Ericson, D.L. Laibson, G. Loewenstien, and J.D. Cohen. Time discounting for primary rewards. *Journal of Neuroscience*, 27(21):5796–5804, 2007.

[19] A. Millner and G. Heal. Collective intertemporal choice: The possibility of time consistency. *NBER Working Paper*, No. 22524, 2016.

[20] D. Pennesi. Uncertain discount and hyperbolic preferences. *Theory and Decision*, pages 1–22, 2015.

[21] J. W. Pratt. Risk aversion in the small and in the large. *Econometrica*, 32(1/2):122–136, 1964.

[22] D. Prelec. Decreasing impatience: A criterion for non-stationary time preference and hyperbolic discounting. *Scandinavian Journal of Economics*, 106(3):511–532, 2004.

[23] F. Proschan. Theoretical explanation of observed decreasing failure rate. *Technometrics*, 5(3):373–383, 1963.

[24] R. T. Rockafellar and R. J.-B. Wets. *Variational analysis*, volume 317 of *A Series of Comprehensive Studies in Mathematics*. Springer-Verlag, Berlin Heidelberg, 1998.

[25] 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.

[26] O. Stein. Twice differentiable characterizations of convexity notions for functions on full dimensional convex sets. *Schedae Informaticae*, 21:55–63, 2012.

[27] K. Takeuchi. Non-parametric test of time consistency: Present bias and future bias. *Games and Economic Behavior*, 71(2):456–478, 2011.

[28] 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.