University of Toronto
Department of Economics

Working Paper 633

The Relation between Behavior under Risk and over Time

By Anujit Chakraborty, Yoram Halevy and Kota Saito

March 31, 2019

The Relation between Behavior under Risk and over Time*

Anujit Chakraborty[†] Yoram Halevy[‡] Kota Saito[§]

February 1, 2019

Abstract

The paper establishes a tight relation between non-standard behaviors in the domains of risk and time, by considering a decision maker with non-expected utility preferences who believes that only present consumption is certain while any future consumption is uncertain. We provide the first complete characterizations of the two-way relations between the certainty effect and present biased temporal behavior, and between the common ratio effect and temporal reversals related to the common difference effect.

JEL: D01, D81, D91

Keywords: time consistency, hyperbolic discounting, non-expected utility, present bias, implicit risk.

---

*This paper subsumes Chakraborty and Halevy's "Allais meets Strotz: Remarks on the Relation between Present Bias and the Certainty Effect" and Saito's "A Relationship between Risk and Time." Discussions with Utpal Sarkar helped in developing the counter-example in Appendix A

[†]University of California, Davis, 1 Shields Ave, CA 95616, USA. E-mail: chakraborty@ucdavis.edu. Web: sites.google.com/site/anujit2006

[‡]Department of Economics, University of Toronto, 150 St. George Street Toronto, ON M5S 3G7 Canada. E-mail: yoram.halevy@utoronto.ca. Web: yoram-halevy.faculty.economics.utoronto.ca. Financial support from SSHRC (FAS # F11-04991) is gratefully acknowledged.

[§]California Institute of Technology, 1200 E California Blvd, MC 228-77 Pasadena, CA 91125, USA. E-mail: saito@caltech.edu. Web: people.hss.caltech.edu/∼saito. Financial support from NSF (SES1558757) is gratefully acknowledged.

1 Introduction

This paper studies if and how behaviors in the domains of risk and time may be similar and related. This similarity is evident in the mutually mirroring mathematical models used for the analysis of behavior under risk and over time. The workhorse model of intertemporal choice, exponential discounting, evaluates the utility of a consumption stream by additively aggregating the utility of each consumption outcome, after exponentially weighting (or discounting) it by the associated time-delay. The canonical model for choice under risk, expected utility, *similarly* calculates the utility of a lottery by aggregating the utility of each possible outcome after weighting it by its respective probability. Further, these normative mathematical models contain similar descriptive inadequacies:

- Preferences are disproportionately sensitive to certainty (certainty effect) in the risk domain and to the present (present bias/ immediacy effect) in the time domain.

- Proportional changes in probabilities (common ratio effect) or the introduction of equal time delays (common difference effect) affect the preferences between two alternatives disproportionately.[1]

Moreover, Keren and Roelofsma (1995) and Weber and Chapman (2005) provide experimental evidence that introducing explicit risk to immediate rewards almost eliminates present bias, while introducing delay to sure outcomes almost eliminates the certainty effect. These parallels are well accepted in the literature (Green and Myerson, 2004; Chapman and Weber, 2006, to name a few) and there is an implicit understanding that the existence of such mirroring behaviors is not a mere coincidence, but points to a common fundamental property of decision making that manifests itself across domains of behavior (Prelec and Loewenstein, 1991; Baucells and Heukamp, 2012). There are a few behavioral channels that have been proposed as explanations for why preference in one domain, may affect preference in the other. For example, a delayed

[1] Often times certainty effect and present bias are taken as special cases of common ratio effect and common difference effect, respectively.

reward or consumption could be inherently risky, as there might be events between the current date and the promised date, which interfere in the process of acquiring the reward (Halevy, 2008). This would explain one direction of the similarity: why risk preferences could influence intertemporal choice patterns. Rachlin et al. (1986; 2000) suggested how the opposite direction of influence could also hold true: if the utility of probabilistic rewards were calculated using mean waiting time before a successful draw of the corresponding reward, then time preferences could be used to derive preferences over the probabilistic rewards. The current paper formalizes these intuitions and provides a two-way characterization of how prominent behavioral traits from the domains of risk and time could be related.

We prove our results in two commonly used decision domains: one where the decision maker (DM) is choosing between temporal rewards ($X \times \mathbb{R}_+$) with the set of time periods being the set of all non-negative numbers and the consumption set is $X \subseteq \mathbb{R}_+$, and another where the DM is choosing from the space of consumption streams ($X^{\mathbb{N}}$, where $\mathbb{N}$ are non-negative integers). In the first domain, we provide a complete relationship between risk and time preferences using two intuitive notions of time-behavior, Temporal Reversals (TR) and Present Biased Temporal Reversals (PBTR), which can be intrinsically linked to hyperbolic discounting and quasi-hyperbolic discounting respectively. We show in Theorem 1 that (i) a decision maker exhibits Strict Common Ratio Effect (SCRE) if and only if his choices satisfy Temporal Reversals (TR), (ii) he exhibits Strict Certainty Effect (SCE) if and only if his choices satisfy Present Biased Temporal Reversals (PBTR), and (iii) he is an expected utility maximizer if and only if he is temporally unbiased (an exponential discounter).

For consumption streams ($X^{\mathbb{N}}$) defined on discrete time, it was previously established in Halevy (2008) and Saito (2011) that Strict Common Ratio Effect (SCRE) was sufficient for Strong Diminishing Impatience (SDI, a property of hyperbolic discounting) and Strict Certainty Effect (SCE) was sufficient for Diminishing Impatience (DI, a property of quasi-hyperbolic discounting) to hold. The property of Diminishing Impatience implies that the ratio of discount weights between two consecutive periods is highest at period zero,

Temporal Rewards ($X \times \mathbb{R}_+$)	Interlink b/t temporal notions across domains	Consump.	Streams	($X^{\mathbb{N}}$)
SCRE $\overset{\text{Theorem 1}}{\Longleftrightarrow}$ TR	TR $\overset{\text{Proposition 3}}{\Longleftrightarrow}$ SDI	SDI	$\overset{\text{Remark 3}}{\Longleftarrow}$ $\overset{\text{Theorem 2}}{\Longrightarrow}$	SCRE WCRE
SCE $\overset{\text{Theorem 1}}{\Longleftrightarrow}$ PBTR	PBTR $\overset{\text{Proposition 3}}{\Longleftrightarrow}$ DIDI	DIDI	$\overset{\text{Remark 3}}{\Longleftarrow}$ $\overset{\text{Theorem 2}}{\Longrightarrow}$	SCE WCE

* Abbreviations used: SCRE=Strict Common Ratio Effect, WCRE=Weak Common Ratio Effect, TR = Temporal Reversals, PBTR = Present Biased Temporal Reversal, SDI = Strong Diminishing Impatience, DIDI = Delay Independent Diminishing Impatience.

Table 1: Summary of our results.

that is, $\dfrac{D(0)}{D(1)} > \dfrac{D(t)}{D(t+1)}$ for the discounted utility $\sum_t D(t)u(c_t)$ of a stream $(c_0, c_1 ...)$. Our main result in the discrete domain is to show that the notion of Diminishing Impatience (DI) does not imply Weak Certainty Effect (WCE) (and hence also does not imply the strict version of Certainty Effect), *unless,* it is adequately extended to hold for all possible delays between streams under consideration. One of our main results shows that, $\dfrac{D(0)}{D(k)} > \dfrac{D(t)}{D(t+k)}$ $\forall k \geq 1$, which we call Delay Independent Diminishing Impatience (DIDI), is the adequate strengthening to derive WCE. We also show that DIDI is a property of commonly used prize-time separable parametric functional forms employed to model present bias. Additionally, a stronger condition, Strong Diminishing Impatience (SDI) implies Weak Common Ratio Effect (WCRE, and also implies Strict Common Ratio Effect for almost all probabilities).[2]

Further, the temporal behaviors considered in the two domains (for e.g, PBTR from $X \times \mathbb{R}_+$ and DIDI from $X^{\mathbb{N}}$), are also interlinked, as shown in Proposition 3. Table 1 succinctly summarizes our results.

The next section provides a brief acknowledgment to the prior literature on risk-time equivalence relations. Sections 3 and 4 provide the relevant definitions from risk and temporal domains of behavior respectively. In Section 5, we state and prove our main results in Theorems 1 and 2. The counter-example that shows the incompleteness of characterization results in the previous lit-

[2]This almost-ness is explained in Corollary 1.

erature (described in Table 2) is included in Appendix A.

2 Background

The idea that Diminishing Impatience (quasi-hyperbolic discounting, present bias) may be related to the certainty of the present and the risk associated with future rewards, was formalized by Halevy (2008). In this model, every consumption path $\mathbf{c} = (c_0, c_1, c_2, \ldots)$ is subject to a constant hazard rate of termination (r), with only the first period of consumption (at $t = 0$) being certain. The DM chooses as if he has the following final utility function on consumption paths: he calculates present discounted utility for every possible length of the path (all periods before termination of consumption). The distribution over present discounted utilities is then evaluated using Rank Dependent Utility (RDU)[3] with probability weighting function $g(\cdot)$. The DM's preferences over consumption streams is represented by the following function:

$$U(\mathbf{c}, r) = \sum_{t=0}^{\infty} g\left((1-r)^t\right) \delta^t u(c_t) \tag{2.1}$$

where δ is a constant pure time preference parameter and $u(\cdot)$ is her felicity function. The DM's *impatience* at time t is the ratio of her discount functions at periods t and $t+1$. Halevy (2008) defines Diminishing Impatience (DI) as the property of "impatience being maximized at $t = 0$". In Table 2 we summarize how the current paper links to Halevy (2008) and Saito (2011).

3 Risky Behavior

In this section, we consider a risk preference $\succsim^r$ on the set of binary lotteries, defined as follows:

$$\Delta = \left\{(x, p; 0, 1-p) \;\middle|\; x \in X \text{ and } p \in [0,1]\right\},$$

[3]For more details of RDU, see Remark 1.

	Results linking risk and time in the Halevy (2008) set-up
Halevy (2008)	DI $\overset{\text{Step 1}}{\Longleftrightarrow}$ Functional inequality* $\underset{\text{Lemma 4.1}}{\overset{\text{Step 2: Segal (1987)}}{\Longleftrightarrow}}$ Increasing elasticity of g(·) $\overset{\text{Step 3}}{\Longleftrightarrow}$ CE
Saito (2011)	(i) Unsubstantiates Segal (1987)'s Lemma 4.1 used in Step 2. (ii) Proposes: DI $\underset{\text{as before}}{\overset{\text{Step 1}}{\Longleftrightarrow}}$ Functional inequality* $\underset{\text{using Steps 2-3}}{\overset{\text{Without}}{\Longleftrightarrow}}$ CE
This paper	Shows DI $\not\Rightarrow$CE & DI $\not\Rightarrow$ Functional inequality* (Appendix A) Shows Halevy (2008), Saito (2011) results are uni-directional CE $\Longrightarrow$ DI. Establishes new two-way results (Theorems 1-2) summarized in Table 1.

*A property of $g(\cdot)$ that is a special case of Kahneman and Tversky (1979, pg.282) subproportionality.

Table 2: Connecting the current paper to previous work.

where X is a non-degenerate closed interval in $\mathbb{R}_+$ including 0. We denote the symmetric part and the asymmetric part by $\sim^r$ and $\succ^r$, respectively. We denote a typical element $(x, p; 0, 1-p)$ of Δ simply by (x, p). We often write $(x, p) \prec^r (y, q)$ when $(y, q) \succ^r (x, p)$.

We formally define the common ratio effect and the certainty effect, which are typical behaviors in the Allais paradox, using the preference $\succsim^r$.

The common ratio effect is characterized as follows: Suppose the subject chooses between a safer option which gives a smaller reward x with a higher probability η, and a riskier option which gives a larger gain y with a lower probability $\eta\mu$, where $\mu < 1$. As η falls, the subject switches his choice from the safe option to the risky option. Formally, the common ratio effect is defined as follows:

Definition 1. $\succsim^r$ is said to exhibit

(i) *Strict Common Ratio Effect (SCRE)*[4] if, for any $x, y \in X$ and $\mu, \tilde{\eta} \in (0, 1]$,

$$(x, \tilde{\eta}) \sim^r (y, \tilde{\eta}\mu) \implies (x, \eta) \prec^r (y, \eta\mu) \text{ for all } \eta \in (0, \tilde{\eta}). \tag{3.1}$$

[4] Under the standard assumptions of monotonicity and continuity axioms, for any $x, y \in X$ and $\tilde{\eta} \in [0, 1]$, there exists μ such that $(x, \tilde{\eta}) \sim^r (y, \tilde{\eta}\mu)$. So the condition cannot be satisfied in a trivial way.

(ii) *Weak Common Ratio Effect (WCRE)* if the conclusion in (3.1) holds with weak preferences and there exist some x, y and $\mu, \tilde{\eta}, \eta$ such that the conclusion in (3.1) holds (with strict preferences).

The general definition provided by Machina (1982, page 305) also becomes equivalent to the above definition within the set of simple binary lotteries. The condition characterizing the Certainty Effect is a special case of the common ratio effect, when $\tilde{\eta} = 1$:

Definition 2. $\succsim^r$ is said to exhibit

(i) *Strict Certainty Effect (SCE)* if, for any $x, y \in X$ and $\mu \in [0, 1)$,

$$(x, 1) \sim^r (y, \mu) \implies (x, \eta) \prec^r (y, \eta\mu) \text{ for all } \eta \in (0, 1). \tag{3.2}$$

(ii) *Weak Certainty Effect (WCE)* if the conclusion of (3.2) holds with weak preferences and there exist some x, y and μ, η such that the conclusion of (3.2) holds (with strict preferences).

By definition, if a DM exhibits Common Ratio Effect, then he exhibits the Certainty Effect. Finally, in the set Δ of binary lotteries, the Independence Axiom reduces to the following:

Definition 3. $\succsim^r$ is said to satisfy *the Independence Axiom* if, for any $x, y \in X$ and $\mu, \eta, \eta' \in [0, 1]$,

$$(x, \eta) \succsim^r (y, \eta\mu) \Leftrightarrow (x, \eta') \succsim^r (y, \eta'\mu).$$

Remark 1. Assume the DM's preferences over binary lotteries are represented by Rank Dependent Utility (RDU), that is $U(x, p) = u(x)\, g(p)$ where $u(\cdot)$ is a real-valued increasing function on X and $g : [0, 1] \to [0, 1]$ is a probability weighting function and for any $\alpha \in \mathbb{R}_+$ there exist $x, y \in X$ such that $\alpha = u(y)/u(x)$. The DM exhibits

(i) SCRE if and only if for all $p, q \in (0, 1)$ and $\ell \in (0, 1]$

$$\frac{g(\ell)}{g(p\ell)} > \frac{g(q\ell)}{g(pq\ell)}. \tag{3.3}$$

(ii) WCRE if and only if (3.3) holds with weak inequality and there exist p, q, ℓ for which (3.3) holds (with strict inequality).

(iii) SCE if and only if $p, q \in (0,1)$

$$g\left(pq\right) > g\left(p\right) g\left(q\right). \tag{3.4}$$

(iv) WCE if and only if (3.4) holds with weak inequality and there exist p, q for which (3.4) holds (with strict inequality).

Proof. We show statement (i). Rest follow similarly.

By assumption, there exist $x, y \in X$ such that $\frac{g(\tilde{\eta})}{g(\tilde{\eta}\mu)} = \frac{u(y)}{u(x)}$. SCRE implies,

$$\forall \eta \in (0, \tilde{\eta}) : \left[\frac{u(y)}{u(x)} > \frac{g(\eta)}{g(\eta\mu)}\right] \ \& \ \forall \eta \in (\tilde{\eta}, 1] : \left[\frac{u(y)}{u(x)} < \frac{g(\eta)}{g(\eta\mu)}\right]. \tag{3.5}$$

(3.5) is equivalent to

$$\forall \eta \in (0, \tilde{\eta}) : \left[\frac{g(\tilde{\eta})}{g(\tilde{\eta}\mu)} > \frac{g(\eta)}{g(\eta\mu)}\right] \ \& \ \forall \eta \in (\tilde{\eta}, 1] : \left[\frac{g(\tilde{\eta})}{g(\tilde{\eta}\mu)} < \frac{g(\eta)}{g(\eta\mu)}\right]. \tag{3.6}$$

Hence, (3.3) implies (3.6). (3.6) implies (3.3) because for any $\eta < \tilde{\eta}$, we can let $q = \eta/\tilde{\eta} < 1$, $\ell = \tilde{\eta}$, and $p = \mu$; for any $\eta > \tilde{\eta}$, we can let $q = \tilde{\eta}/\eta < 1$, $\ell = \eta$, and $p = \mu$. □

4 Intertemporal Behavior

In this section, we define preferences that subsume the classes of exponential, hyperbolic and quasi-hyperbolic discounting. We denote the set of time periods by T. In the following, we consider two cases. First, when preferences are defined over temporal rewards ($X \times T$ when $T = \mathbb{R}_+$), and then over consumption streams (X^T when $T = \mathbb{N}$).

4.1 Temporal rewards in continuous time

For each $d \in T$, we denote the set of temporal rewards, paid after time d, by $X(d) = \{[x,t] \mid x \in X \text{ and } t \in T \text{ such that } t \geq d\}$. The DM's time-indexed preferences are given by $\{\succsim_d\}_{d\in T}$, where $\succsim_d$ is a binary relation on $X(d)$ for each decision time $d \in T$.[5]

Hyperbolic discounting describes the following pattern of dynamic choice: the subject chooses a later-larger reward over a earlier-smaller reward, but he reverses his choice as both reward dates approach the decision date.[6] Temporal Reversal formalizes this behavioral pattern as follows:[7]

Definition 4. $\{\succsim_d\}_{d\in T}$ is said to exhibit:

(i) *Temporal Reversal (TR)* if, for any $x, y \in X$ and $\tilde{d}, t, s \in T$ such that $\tilde{d} \leq t \leq s$,

$$[x,t] \sim_{\tilde{d}} [y,s] \implies \begin{cases} [x,t] \prec_d [y,s] \text{ for all } d \text{ such that } d < \tilde{d}, \\ [x,t] \succ_d [y,s] \text{ for all } d \text{ such that } t > d > \tilde{d}. \end{cases}$$

(ii) *Present Biased Temporal Reversal (PBTR)* if, for any $x, y \in X$ and $t, s \in T$ such that $t \leq s$,

$$[x,t] \sim_t [y,s] \implies [x,t] \prec_d [y,s] \text{ for all } d < t.$$

(iii) *Temporally Unbiased (TU)* if, for any $x, y \in X$ and $d, d', s, t \in \mathbb{R}_+$,

$$[x,t] \succsim_d [y,s] \Leftrightarrow [x,t] \succsim_{d'} [y,s].$$

[5]For each $d \in T$, we denote the symmetric part and the asymmetric part of $\succsim_d$ by $\sim_d$ and $\succ_d$, respectively.

[6]In the following three definitions of time preferences, we focus on positive payoffs for simplicity. For the case of negative payoffs, present bias appears as procrastination and is defined in the same way by switching strict preference from $\succ$ to $\prec$, and vice versa. O'Donoghue and Rabin (1999) offers examples of procrastination and Halevy (2008, page 1157) discusses how to incorporate into the current framework using the reflection effect.

[7]Note that the characterization of hyperbolic discounting in Proposition 1 of Dasgupta and Maskin (2005, page 1293) is exactly the same as above.

PBTR is a special case of TR when $\tilde{d} = t$. By definition, if a DM exhibits TR, then he exhibits PBTR behavior, but the converse is not true. Temporal reversals, or lack thereof, are defined here in terms of dynamic decision making, thus assuming time invariance (Halevy, 2015). TU preferences are time consistent and correspond to exponential discounting.

4.2 Consumption streams in discrete time

We now consider consumption streams in discrete time (i.e., $T = \mathbb{N}$). In this case, temporal behavior is usually characterized by properties of the discount-function, and under time invariance it depends only on the distance between the evaluation time and consumption time. Let $D(\cdot)$ be the DM's discount-function, so the utility of consuming x after τ periods is $D(\tau)u(x)$, where u is a real valued function on X. D exhibits *hyperbolic discounting* (HD) if $D(\tau) = 1/(1 + \rho\tau)$ for some $\rho > 0$; is *quasi-hyperbolic discounting* (QHD) if $D(0) = 1$ and $D(\tau) = \beta\delta^\tau$ for some $\delta \in (0, 1]$ and $\beta < 1$ for all $\tau \geq 1$.

The DM's (one period) impatience at t is $D(t)/D(t+1)$. DM's (k period) impatience at t is $D(t)/D(t+k)$. In Table 3 we define the notions of temporal behavior for streams ($X^{\mathbb{N}}$).

	Definition
Diminishing Impatience (DI)	$\frac{D(0)}{D(1)} > \frac{D(t)}{D(t+1)} \; \forall t \in \mathbb{N}_+$
Delay Independent Diminishing Impatience (DIDI)	$\frac{D(0)}{D(k)} > \frac{D(t)}{D(t+k)} \; \forall k, t \in \mathbb{N}_+$
Strongly Diminishing Impatience (SDI)	$\frac{D(t)}{D(t+1)} > \frac{D(t')}{D(t'+1)} \; \forall t, t' \in \mathbb{N}$ with $t < t'$
Delay Independent Strongly Diminishing Impatience (DISDI)	$\frac{D(t)}{D(t+k)} > \frac{D(t')}{D(t'+k)} \; \forall t, t' \in \mathbb{N}, k \in \mathbb{N}_+$ with $t < t'$

Table 3: Notions of temporal behavior for consumption streams ($X^{\mathbb{N}}$).

Proposition 1.

(i) DIDI implies DI (but DI does not imply DIDI).

(ii) SDI and DISDI are equivalent.

Proof. (i) DI is the special case of DIDI where delay $k = 1$. An implication of the counter-example provided in Appendix A is that DI does not imply DIDI.

(ii) DISDI trivially implies SDI. Assume SDI and to show DISDI, fix k, t', t such that $t' > t$ to show that $\frac{D(t)}{D(t+k)} > \frac{D(t')}{D(t'+k)}$. Notice that

$$\frac{D(t)}{D(t+k)} = \underbrace{\frac{D(t)}{D(t+1)} \frac{D(t+1)}{D(t+2)} \cdots \frac{D(t+k-1)}{D(t+k)}}_{k-\text{terms}} = \prod_{d=0}^{k-1} \frac{D(t+d)}{D(t+d+1)},$$

$$\frac{D(t')}{D(t'+k)} = \underbrace{\frac{D(t')}{D(t'+1)} \frac{D(t'+1)}{D(t'+2)} \cdots \frac{D(t'+k-1)}{D(t'+k)}}_{k-\text{terms}} = \prod_{d=0}^{k-1} \frac{D(t'+d)}{D(t'+d+1)}.$$

Since $t' > t$, by SDI, we have for each $d \in \{0, \ldots, k-1\}$, $\frac{D(t+d)}{D(t+d+1)} > \frac{D(t'+d)}{D(t'+d+1)}$. Hence, $\frac{D(t)}{D(t+k)} > \frac{D(t')}{D(t'+k)}$. □

DI and SDI have been proposed by Halevy (2008). DIDI and DISDI are new properties motivated by the hyperbolic-discounting and the quasi-hyperbolic discounting models, and they focus on the failure of stationarity independently of the delay under consideration.[8] DIDI requires impatience to diminish for all possible delays ($k \geq 1$), hence is a strengthening of DI. Proposition 2 shows that the standard behavioral models satisfy DIDI and DISDI.

Proposition 2.

(i) QHD satisfies DIDI but not DISDI.

(ii) HD satisfies DISDI (and hence SDI and DIDI).

Proof. (i) For $t' > t > 0$

$$\frac{D(0)}{D(k)} = \frac{1}{\beta\delta^k} > \frac{1}{\delta^k} = \frac{D(t)}{D(t+k)} = \frac{D(t')}{D(t'+k)}$$

(ii) For arbitrary k, and $t' > t \geq 0$,

$$\frac{D(t)}{D(t+k)} = 1 + \frac{\rho k}{1+\rho t} > 1 + \frac{\rho k}{1+\rho t'} = \frac{D(t')}{D(t'+k)}$$

□

Finally, we elaborate on the relationship between the definitions for temporal rewards in continuous time (Definition 4) and the definitions for discrete time consumption streams (Definitions in Table 3). To achieve this we consider the intersection of their domains, which is temporal rewards in discrete time.

Proposition 3. *Suppose that $T = \mathbb{N}$ and there exist $D : T \to [0,1]$ and $u : X \to \mathbb{R}_+$ such that for each $d \in T$, $\succsim_d$ is represented by $U_d([x,t]) = D(t-d)u(x)$ and for any $\alpha \in \mathbb{R}_+$ there exist $x, y \in X$ such that $\alpha = u(x)/u(y)$. Then, the following results hold:*

(i) D exhibits SDI (and hence DISDI) if and only if $\{\succsim_d\}_{d\in T}$ exhibits TR.

(ii) D exhibits DIDI if and only if $\{\succsim_d\}_{d\in T}$ exhibits PBTR.

[8] Halevy (2015) provides a formal definition and recent experimental evidence for stationarity in a dynamic setting.

Proof. For this proof, we will use that SDI and DISDI are equivalent. To show (i), suppose that $\{\succsim_d\}$ exhibit TR and show that D exhibits DISDI, i.e, for an arbitrary non negative integer $\tau' < \tau$ we will show that

$$\frac{D(\tau')}{D(\tau'+k)} < \frac{D(\tau)}{D(\tau+k)}.$$

Choose $x, y \in X$ such that $D(\tau')/D(\tau'+k) = u(y)/u(x)$. Hence, $[x,\tau] \sim_{\tau-\tau'} [y,\tau+k]$. Then by TR we have $[x,\tau] \prec_0 [y,\tau+k]$. This means that $D(\tau')u(x) > D(\tau'+k)u(y)$. It follows that, $\dfrac{D(\tau)}{D(\tau+k)} = \dfrac{u(y)}{u(x)} < \dfrac{D(\tau')}{D(\tau'+k)}$. Since the choice of τ is arbitrary, this means that D exhibits SDI.

Converse direction: Suppose that there exist $x, y \in X$ and $\tilde{d}, t, s \in T$ such that $[x,t] \sim_{\tilde{d}} [y,s]$ and $\tilde{d} \leq t \leq s$. Choose $d, d' \in T$ such that $d < \tilde{d}$ and $d' > \tilde{d}$ to show $[x,t] \prec_d [y,s]$ and $[x,t] \succ_{d'} [y,s]$.

Since $[x,t] \sim_{\tilde{d}} [y,s]$, by definition, $D(t-\tilde{d})u(x) = D(s-\tilde{d})u(y)$, so that

$$\frac{D(t-\tilde{d})}{D(s-\tilde{d})} = \frac{u(y)}{u(x)}. \tag{4.1}$$

Since $d < \tilde{d} < d'$, it follows from DISDI that

$$\frac{D(t-d)}{D(s-d)} < \frac{D(t-\tilde{d})}{D(s-\tilde{d})} < \frac{D(t-d')}{D(s-d')}.$$

By (4.1), we get $D(t-d)u(x) < D(s-d)u(y)$ and $D(t-d')u(x) > D(s-d')u(y)$. Hence, $[x,t] \prec_d [y,s]$ and $[x,t] \succ_{d'} [y,s]$.

To show (ii) take $\tau' = 0$ in the forward direction to get DIDI from PBTR. Use $t = \tilde{d}$ and consider $d < \tilde{d}$ for the converse . □

5 Results

5.1 Results for temporal rewards in continuous time

In this subsection, we assume $T = \mathbb{R}_+$. We assume that the DM's temporal and risk preferences are connected in the following way: the individual discounts future rewards because he is uncertain whether he can consume it. We model this uncertainty through a stopping process that determines the last period until which rewards are available. Let $p(t)$ be the probability that the DM may collect a reward at time t. At time d, such that $0 \leq d \leq t$, the DM updates the probability that a reward is available at time t according to Bayes rule: $p(t|d) = p(t)/p(d)$. Therefore, at time d he prefers receiving the temporal reward $[x, t]$, to another reward $[y, s]$ if and only if his risk preferences rank the binary lottery $(x, p(t|d))$ (which gives x with the probability $p(t|d)$) over the lottery $(y, p(s|d))$. Thus, the DM's time preferences $\{\succsim_d\}_{d\in T}$ for each decision time $d \in T$ and risk preferences $\succsim^r$ are related as follows:
Assumption 1: For all $d \in T$ and $[x, t], [y, s] \in X(d)$,

$$[x, t] \succsim_d [y, s] \Leftrightarrow \big(x, p(t|d)\big) \succsim^r \big(y, p(s|d)\big). \tag{5.1}$$

We additionally assume a regularity condition on future uncertainty, which states that immediate rewards are certain, but as the promised date for future rewards becomes increasingly distant, the probability of receiving the reward continuously decreases to zero:
Assumption 2: $p(0) = 1$, $p(\infty) = 0$, and $p(\cdot)$ is strictly decreasing and continuous.

Theorem 1. *Suppose that Assumptions 1 and 2 hold.*

(i) $\succsim^r$ *exhibits SCRE iff* $\{\succsim_d\}_{d\in T}$ *exhibits TR.*

(ii) $\succsim^r$ *exhibits SCE iff* $\{\succsim_d\}_{d\in T}$ *exhibits PBTR.*

(iii) $\succsim^r$ *satisfies the Independence Axiom iff* $\{\succsim_d\}_{d\in T}$ *is TU.*

Proof. To prove (i), suppose that $\succsim^r$ exhibits SCRE. Choose any $x, y \in X$ and $\tilde{d}, t, s \in T$ such that $[x,t] \sim_{\tilde{d}} [y,s]$ and $\tilde{d} \leq t \leq s$. Then by definition, $(x, p(t|\tilde{d})) \sim^r (y, p(s|\tilde{d})) = (y, p(s|t)p(t|\tilde{d}))$. Fix $d < \tilde{d}$ to show $[x,t] \prec_d [y,s]$. Since p is strictly decreasing, $p(d) > p(\tilde{d})$, so that $p(t|d) < p(t|\tilde{d})$. So SCRE implies that $(x, p(t|d)) \prec^r (y, p(s|t)p(t|d)) = (y, p(s|d))$. Then by definition, $[x,t] \prec_d [y,s]$.

To show the converse, suppose that $\{\succsim_d\}$ exhibits TR. Choose any $x, y \in X$ and $\mu, \tilde{\eta} \in [0,1]$ such that $(x, \tilde{\eta}) \sim^r (y, \tilde{\eta}\mu)$. Fix $\eta \in (0, \tilde{\eta})$ to show $(x, \eta) \prec^r (y, \eta\mu)$. Since p is a strictly decreasing bijection to $[0,1]$, there exist t and $\tilde{d}$ such that $t \geq \tilde{d} > 0$ and $p(t) = \eta$ and $p(\tilde{d}) = \eta/\tilde{\eta}$. Then, $p(t|\tilde{d}) = \tilde{\eta}$. Also, there exists s such that $s \geq t$ and $p(s) = \mu\eta$. Then, $p(s|t) = \mu$. Hence, $(x, p(t|\tilde{d})) \sim^r (y, p(s|t)p(t|\tilde{d})) = (y, p(s|\tilde{d}))$, so that $[x,t] \sim_{\tilde{d}} [y,s]$, by definition. Therefore, by TR, $[x,t] \prec_0 [y,s]$. So the definition shows that $(x, \eta) = (x, p(t)) \prec^r (y, p(s)) = (y, \eta\mu)$.

The proof for part (ii) is very similar to part (i) and is hence omitted.

To show (iii). Suppose that $\succsim^r$ satisfies the Independence Axiom. Choose any $x, y \in X$ and $t, s, d, d' \in T$ such that $[x,t] \succsim_d [y,s]$ to show $[x,t] \succsim_{d'} [y,s]$. Since $[x,t] \succsim_d [y,s]$, by definition $(x, p(t|d)) \succsim^r (y, p(s|d))$. Consider the case in which $d > d'$. By the Independence Axiom, $(x, p(t|d')) = (x, p(t|d)p(d|d')) \succsim^r (y, p(s|d)p(d|d')) = (y, p(s|d'))$. By the definition, $[x,t] \succsim_{d'} [y,s]$. The proof for the other case in which $d' > d$ is similar.[9]

To show the converse, suppose that $\{\succsim_d\}$ is TU. Choose any $x, y \in X$ and $\mu, \eta, \eta' \in [0,1]$ such that $(x, \eta) \succsim^r (y, \eta\mu)$ to show $(x, \eta') \succsim^r (y, \eta'\mu)$. Consider the case where $\eta' > \eta$. Since p is strictly decreasing and bijection to $[0,1]$, there exist $t, s, d \in T$ such that $s \geq t$, $p(t) = \eta$, $p(s) = \eta\mu$, and $p(d) = \eta/\eta'$. Then, $p(t|d) = \eta'$ and $p(s|d) = \eta'\mu$. Since $\{\succsim_d\}$ is TU, $(x, \eta) \succsim^r (y, \eta\mu) \leftrightarrow [x,t] \succsim_0 [y,s] \leftrightarrow [x,t] \succsim_d [y,s] \leftrightarrow (x, p(t|d)) \succsim^r (y, p(s|d)) \leftrightarrow (x, \eta') \succsim^r (y, \eta'\mu)$. The proof for the other case in which $\eta' < \eta$ is similar.[10] □

[9] Since $(x, p(t|d)) = (x, p(t|d')p(d'|d)) \succsim^r (y, p(s|d')p(d'|d)) = (y, p(s|d))$. Since $(x, p(t|d)) \succsim^r (y, p(s|d))$, by the Independence Axiom, $(x, p(t|d')) \succsim^r (y, p(s|d'))$. Hence, $[x,t] \succsim_{d'} [y,s]$.

[10] Since p is strictly decreasing and bijection to $[0,1]$, there exist $t, s, d \in T$ such that $s \geq t$, $p(t) = \eta'$, $p(s) = \eta'\mu$, and $p(d) = \eta'/\eta$. Then, $p(t|d) = \eta$ and $p(s|d) = \eta\mu$. Since $\{\succsim_d\}$

Part (iii) of the theorem has an important implication: the approach of assuming a non-constant or uncertain hazard function combined with standard expected utility, used to explain non-stationary choices in psychology and biology (Kagel et al. (1986), Green and Myerson (2004),Sozou (1998)), cannot explain time inconsistency. This is similar to the findings discussed in Halevy (2004; 2015).

The proof of Theorem 1 relies on the structural similarity between risky and intertemporal choices: a decrease in the risk is equivalent to the time of the reward and the decision time getting closer. Such a similarity had been also suggested by Prelec and Loewenstein (1991), although they did not provide a formal argument. The Probability Time Tradeoff axiom proposed by Baucells and Heukamp (2012) and used in Chakraborty (2016) to axiomatize preferences on a richer domain of intertemporal lotteries has a similar flavor.

5.2 Results for consumption streams in discrete time

Given the representation (2.1), the composite discount function at period t is:

$$D\left(t\right)=\delta^{t}g\left(\left(1-r\right)^{t}\right). \tag{5.2}$$

Remark 2. Consider a DM represented by (2.1) with continuous $g\left(\cdot\right)$.

(i) DI holds if and only if for every $r\in(0,1)$ and $t\in\mathbb{N}_{+}$:

$$g\left(\left(1-r\right)^{t+1}\right)>g\left(\left(1-r\right)\right)g\left(\left(1-r\right)^{t}\right). \tag{5.3}$$

(ii) DIDI holds if and only if for every $r\in(0,1)$ and $t,k\in\mathbb{N}_{+}$:

$$g\left(\left(1-r\right)^{t+k}\right)>g\left(\left(1-r\right)^{k}\right)g\left(\left(1-r\right)^{t}\right). \tag{5.4}$$

is TU, $(x,\eta)\succsim^{r}(y,\eta\mu)\leftrightarrow(x,p(t|d))\succsim^{r}(y,p(s|d))\leftrightarrow[x,t]\succsim_{d}[y,s]\leftrightarrow[x,t]\succsim_{0}[y,s]\leftrightarrow(x,p(t))\succsim^{r}(y,p(s))\leftrightarrow(x,\eta')\succsim^{r}(y,\eta'\mu)$.

(iii) SDI holds if and only if for every $r \in (0,1)$ and $t, t' \in \mathbb{N}$ such that $t < t'$:

$$\frac{g\left((1-r)^{t}\right)}{g\left((1-r)^{t+1}\right)} > \frac{g\left((1-r)^{t'}\right)}{g\left((1-r)^{t'+1}\right)}. \tag{5.5}$$

(iv) DISDI holds if and only if for every $r \in (0,1)$, $t, t' \in \mathbb{N}$, $k \in \mathbb{N}_+$ such that $t < t'$:

$$\frac{g\left((1-r)^{t}\right)}{g\left((1-r)^{t+k}\right)} > \frac{g\left((1-r)^{t'}\right)}{g\left((1-r)^{t'+k}\right)}. \tag{5.6}$$

The proofs follow immediately from the definitions. Next, we summarize the implications of risk attitude on the intertemporal preferences in (2.1).

Remark 3. Consider a DM represented by (2.1) with continuous $g(\cdot)$

(i) SCE (3.4) implies DIDI (i.e., (5.4))

(ii) SCRE (3.3) implies SDI (i.e., (5.5)) (and, hence, DISDI (i.e., (5.6))).

The first claim holds by letting $p = (1-r)^k$ and $q = (1-r)^t$; the second claim holds by letting $p = 1-r$, $q = (1-r)^{t'-t}$, and $\ell = (1-r)^t$.

For the relation in the direction from time to risk, DI as defined above does *not* imply Weak Certainty Effect for general weighting functions. The certainty effect implies a bias towards certainty irrespective of how risky the alternative is, the dual to which would be a bias towards the present ($t = 0$) irrespective of the delay between the two prospects being compared. In evaluating the reason for the severed connection between time and risk preferences, we note that the definition of diminishing impatience used in the literature focuses on a delay of a single period, thus only comparing $D(t)$ to $D(t+1)$ as t increases from 0. This one-period definition fails to generalize to longer delays, as shown in our counter-example in Appendix A, and thus fails to account for present bias behaviorally. Theorem 2 below shows that DIDI is sufficient for WCE, and SDI is sufficient for WCRE.

Theorem 2. *Consider a DM represented by (2.1) with continuous* $g(\cdot)$.

(i) SDI implies WCRE.

(ii) DIDI implies WCE.

Proof. We first show claim (i). Assume SDI. By Remarks 1 and 2, it suffices to show that (5.5) in Remark 2 implies that (3.3) in Remark 1 holds with weak inequality and there exist p, q, ℓ for which (3.3) holds. Let $p = 1 - r$, $q = (1 - r)^{t' - t}$, and $\ell = (1 - r)^t$. Then (3.3) holds.

Now we will show that for any $p, q \in (0, 1)$ and $\ell \in (0, 1]$, (3.3) in Remark 1 holds with weak inequality. Since SDI and DISDI are equivalent, in the following, we assume DISDI. Consider a sequence $\{\frac{m_i}{n_i}\}_{i=1}^{\infty}$ of rational numbers that converges to $\frac{\ln p}{\ln q\ell}$, where m_i, n_i are positive integers. Similarly, consider a sequence $\{\frac{a_i}{b_i}\}_{i=1}^{\infty}$ of positive rational numbers that converges to $\frac{\ln \ell}{\ln q\ell}$, where a_i, b_i are positive integers. Note that $\frac{\ln \ell}{\ln q\ell} < 1$, so we can choose $\{\frac{a_i}{b_i}\}_{i=1}^{\infty}$ such that $a_i < b_i$. Now, given this sequences, define a sequence $\{r_i\}$, such that $q\ell = r_i^{n_i b_i}$, that is $r_i = (q\ell)^{\frac{1}{n_i b_i}} < 1$. Note that as $\frac{a_i}{b_i}$ converges to $\frac{\ln \ell}{\ln q\ell}$, $r_i^{a_i n_i} = (q\ell)^{\frac{a_i}{b_i}}$ converges to $(q\ell)^{\frac{\ln \ell}{\ln q\ell}} = \ell$. Similarly, as $\frac{m_i}{n_i}$ converges to $\frac{\ln p}{\ln q\ell}$, $r_i^{m_i b_i} = (q\ell)^{\frac{m_i}{n_i}}$ converges to $(q\ell)^{\frac{\ln p}{\ln q\ell}} = p$.

Now using DISDI, $\forall i$:

$$\frac{g\left(r_i^{a_i n_i}\right)}{g\left(r_i^{a_i n_i + m_i b_i}\right)} > \frac{g\left(r_i^{n_i b_i}\right)}{g\left(r_i^{n_i b_i + m_i b_i}\right)}$$

Using the continuity of g, as $i \to \infty$, WCRE follows:

$$\frac{g(\ell)}{g(p\ell)} \geq \frac{g(q\ell)}{g(pq\ell)}$$

Next, we will show (ii). Assume DIDI. By Remarks 1 and 2, it suffices to show that (5.6) in Remark 2 implies that (3.4) in Remark 1 holds with weak inequality and there exist p, q for which (3.4) holds. Let $p = (1 - r)^k$ and $q = (1 - r)^t$. Then (3.4) holds.

Part (ii) is a special case of (i), where $\ell = 1$, $a_i = 0$, $b_i = 0$, and DIDI replaces DISDI. $\square$

In the above theorem, we have WCRE and WCE as behavioral implications, but not the versions with strict inequalities (SCRE and SCE). This gap is inevitable given the difference between the connectedness in the domain of risk preferences (i.e, the probabilities are numbers in $[0,1]$) and the non-connectedness of the domain of time preferences (i.e., the dates are non-negative integers).[11] The last step of the proof is to approximate a real number by the limit of rational numbers. When we take the limit, the related strict inequality inherited from behavior in the time domain becomes a weak inequality. We show in Corollary 1, that SDI implies SCRE for almost all probabilities; DI implies SCE for almost all probabilities.

Corollary 1. *Consider a DM represented by (2.1) with continuous $g(\cdot)$.*

(i) There exists a dense subset Δ_1 of $(0,1)^2 \times (0,1]$ such that SDI implies for any $(p,q,\ell) \in \Delta_1$

$$\frac{g(\ell)}{g(p\ell)} > \frac{g(q\ell)}{g(pq\ell)}.$$

(ii) There exists a dense subset Δ_2 of $(0,1)^2$ such that DIDI implies for any $(p,q) \in \Delta_2$

$$g(pq) > g(p)g(q).$$

Proof. For (i), define $\Delta_1 = \{(r^k, r^s, r^t) | r \in (0,1), k,s \in \mathbb{N}_+, t \in \mathbb{N}\}$. Notice that $(r_i^{m_i b_i}, r_i^{(b_i - a_i)n_i}, r_i^{a_i n_i})$ in the proof of Theorem 2 is a sequence in Δ_1 that converges to $(p,q,\ell) \in (0,1)^2 \times (0,1]$. Part (ii) is proved similarly. □

[11] The proof of Theorem 1 in continuous time makes it clear that the complete relation, especially the relation from time preferences to risk preferences relies on the continuous time structure.

A Diminishing Impatience does not imply Weak Certainty Effect

The following counter-example proves that DI does not imply Weak Certainty Effect (and hence by Theorem 2 also does not imply DIDI). If (5.3) implied (3.4), then (5.3) would also imply that $\forall r \in (0,1)$ and $\forall m, n \in \mathbb{N}$

$$g(r^{m+n}) \geq g(r^m)g(r^n) \tag{A.1}$$

We will rewrite these expressions in an additive form by defining $f(x) = -log(g(e^{-x})) \iff g(x) = e^{-f(-logx)}$. Then $f : (0,\infty) \to (0,\infty)$ is differentiable and increasing, just like the function g. The inequalities under consideration are now:

$$\begin{aligned} \forall t \in \mathbb{N} \text{ and } \forall r \in (0,1),\ g(r^{t+1}) &> g(r)g(r^t) \\ \iff e^{-f\left(-log\left(r^{t+1}\right)\right)} &> e^{-f\left(-log\left(r^t\right)\right)}e^{-f(-log(r))} \\ \iff f(-(t+1)log(r)) &< f(-tlog\,(r)) + f(-log(r)) \end{aligned}$$

Now, defining $x := -log\,(r) \in (0,\infty)$ for $r \in (0,1)$.

$$f((t+1)x) < f(tx) + f(x) \tag{A.2}$$

Further, the boundary conditions $g(0) = 0$ and $g(1) = 1$ translate to $f(0) = 0$ and $f(\infty) = \infty$.[12]
Similarly, (A.1) converts to

$$f(mx + nx) \leq f(mx) + f(nx) \quad \forall x \in (0,\infty) \text{ and } \forall m, n \in \mathbb{N} \tag{A.3}$$

Summing it up, (5.3) implies (A.1), if and only if (A.2) implies (A.3). The next step is to propose a function f which would satisfy (A.2) on all points of

[12] Using the extended real line $(\mathbb{R} \cup \infty)$

its domain, but, for some $x \in \mathbb{R}$ and some $m, n \in \mathbb{N}$,

$$f(mx + nx) > f(mx) + f(nx) \tag{A.4}$$

Instead of providing the function f, we propose it's derivative h, so f can be calculated as $f(x) = \int_0^x h(x)dx$.[13] Let, $k = \frac{20}{1+\sin(\pi/2-.0001)}$ and $\delta = 50k\pi\cos(\pi/2 - .0001) \approx .157$.

Let,

$$h(x) = \begin{cases} 11 + (1-x)\delta & For\ x < 1 \\ 1 + \frac{k}{2} + \frac{k}{2}\sin 100\pi(1 + \frac{\pi/2-.0001}{100\pi} - x) & For\ 1 \leq x \leq 1.005 + \frac{\pi/2-.0001}{100\pi} \\ 1 & For\ 1.005 + \frac{\pi/2-.0001}{100\pi} < x < 2 - .005 \\ 4 + 3\sin 100\pi(x-2) & For\ 2 - .005 \leq x \leq 2 + .005 \\ 7 & For\ 2 + .005 < x < 2.5 - .005 \\ 4 + 3\sin 100\pi(2.5 - x) & For\ 2.5 - .005 \leq x \leq 2.5 + .005 \\ 1 & For\ 2.5 + .005 < x < 3 - .005 \\ 4 + 3\sin 100\pi(x-3) & For\ 3 - .005 \leq x \leq 3 + .005 \\ 7 & For\ 3 + .005 < x < 5 - .005 \\ 4 + 3\sin 100\pi(5 - x) & For\ 5 - .005 \leq x \leq 5 + .005 \\ 1 & For\ x > 5 + .005 \end{cases}$$

f is increasing, twice differentiable and $f(\infty) = \infty$. $h(x)$ is plotted in Figure A.1.

We next show that (A.2) holds.

Lemma 1. $\forall t \in \mathbb{N},\ \forall x \in \mathbb{R}, \int_0^x h(x)dx > \int_{tx}^{(t+1)x} h(x)dx$.

Proof. The most intuitive way to check the claim would be to notice that the sinusoids introduced hardly perturb the area under the curve. Figure A.2 illustrates the point in a more clear fashion by considering the function

[13] Recall that $f(0) = 0$.

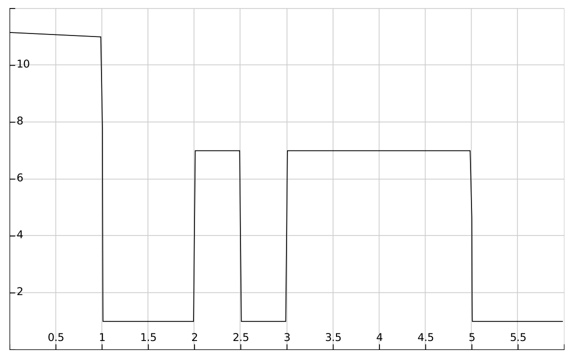

Figure A.1: The function h.

h for a small part of the real line. For all practical purposes, one could go about checking the inequalities by replacing the sinusoid (in black) in Figure 1 by a corresponding discontinuous function($\bar{h}(x) = 7$ for $x \leq 2.5$, $\bar{h}(x) = 1$ for $x > 2.5$ as drawn in red). The area between the two curves in $[2.495, 2.5]$ is only $(.005 * 3 - \frac{3}{100\pi}) \approx .005$. Therefore, as long as the inequalities hold with a large enough margin, this intuitive method of approximating sinusoids with flat lines works fine. The area between the two curves in $[2.5, 2.505]$ is also $(.005 * 3 - \frac{3}{100\pi})$. Thus, the two approximations in $[2.495, 2.505]$ are equal and opposite in direction, and the areas under the red and black curves in this region are equal. During our analysis, in some cases there will be multiple approximations in opposite directions which would partially or completely cancel each other out.

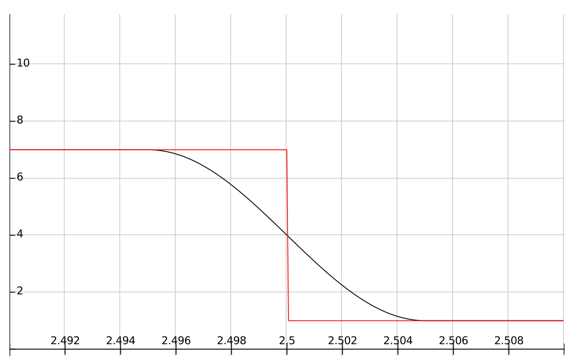

Figure A.2: Function h approximated in a sinusoidal region

Utilizing this intuition more rigorously, one can create upper bounds and lower bounds on $\int_{tx}^{(t+1)x} h(x)dx$ and $\int_0^x h(x)dx$ respectively to complete the proof. For $0 < x \leq 1$, $\int_0^x h(x)dx > \int_{tx}^{(t+1)x} h(x)dx$ is obvious, as $[0, x]$ contains the highest values obtained by $h(x)$ on the real line.
For, $1 < x \leq \frac{5}{3}$, $\int_0^x h(x)dx = \int_0^1 h(x)dx + \int_1^x h(x)dx > \frac{1}{2}(11+11+\delta)+(x-1) = 10 + \frac{\delta}{2} + x$.[14] The inequality holds because $h(x) \geq 1$ with strict inequality for $1 \leq x < 1.005 + \frac{\pi/2-.0001}{100\pi}$, and hence $\int_1^x h(x)dx > x - 1$. In the interval $[tx, (t+1)x]$, $h(x) \leq 7$ and after mutual canceling out there are no more than 3 sinusoidal perturbations which could increase the area under the curve. Hence, $\int_{tx}^{(t+1)x} h(x)dx < 7x + 3(.015 - \frac{3}{100\pi}) = 6x + x + 3(.015 - \frac{3}{100\pi}) \leq 6(\frac{5}{3}) + x + 3(.015 - \frac{3}{100\pi}) = 10 + x + 3(.015 - \frac{3}{100\pi})$.
For $\frac{5}{3} \leq x \leq 2$, $\int_0^x h(x)dx > 10 + \frac{\delta}{2} + x$ as before. On the other hand, using the same line of logic as before, $\int_x^{2x} g(x)dx < 1.x + 6[(4-3) + (2.5 - 2)] + 3.(.015 - \frac{3}{100\pi}) = 9 + x + 3.(.015 - \frac{3}{100\pi})$. Similarly, $\int_{2x}^{3x} h(x)dx \leq 1.x + 6[5 - 2.\frac{5}{3}] + 3.(.015 - \frac{3}{100\pi}) = 10 + x + 3.(.015 - \frac{3}{100\pi})$.
Similarly for larger values of x, it can be shown that $\int_0^x h(x)dx > \int_{tx}^{(t+1)x} h(x)dx$. (follows trivially for $x \geq 5$.) □

Now complete the counter-example:

$$\int_0^2 h(x)dx < 12+\frac{\delta}{2}+\{.01*10+(.015-\frac{3}{100\pi})\} < 14-2(.015-\frac{3}{100\pi}) = \int_3^5 h(x)dx$$

The first inequality follows from setting an upper bound on the sinusoidal perturbation introduced around 1.[15] Therefore, $f(5) > f(2) + f(3)$, which provides us with the counter-example to equation (A.3) and hence, to equation (A.1). In other words, as (A.2) does not imply (A.3), (5.3) does not imply (A.1), and hence, (5.3) does not imply (3.4).
That is, even if for all $t \in \mathbb{N}$ and for all $r \in (0,1)$: $g((1-r)^{t+1}) > g((1-r)^t)g((1-r))$ it does not imply that $\forall p, q \in (0,1)$: $g(pq) \geqslant g(p)g(q)$.

[14] $\delta = 50k\pi\cos(\pi/2 - .0001) = .157$ (approximately)

[15] This particular sinusoid dies down after $1.005 + \frac{\pi/2-.0001}{100\pi} < 1.01$ and never rises above the $h(x) = 1$ line by more than 6 units.

References

Baucells, Manel and Franz H. Heukamp, "Probability and Time Trade-Off," *Management Science*, April 2012, *58* (4), 831–842.

Chakraborty, Anujit, "Present Bias," 2016.

Chapman, Gretchen B and Bethany J Weber, "Decision biases in intertemporal choice and choice under uncertainty: testing a common account," *Memory & Cognition*, 2006, *34* (3), 589–602.

Dasgupta, Partha and Eric Maskin, "Uncertainty and hyperbolic discounting," *The American Economic Review*, 2005, *95* (4), 1290–1299.

Green, Leonard and Joel Myerson, "A Discounting Framework for Choice with Delayed and Probabilistic Rewards," *Psuchological Bulletin*, 2004, *130*, 769–792.

Halevy, Yoram, "Diminishing impatience: disentangling time preference from uncertain lifetime," 2004.

_ , "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect," *American Economic Review*, 2008, *98* (3), pp. 1145–1162.

_ , "Time Consistency: Stationarity and Time Invariance," *Econometrica*, January 2015, *83* (1), 335–352.

Kagel, John H, Leonard Green, and Thomas Caraco, "When foragers discount the future: constraint or adaptation?," *Animal Behaviour*, 1986, *34*, 271–283.

Keren, Gideon and Peter Roelofsma, "Immediacy and Certainty in Intertemporal Choice," *Organizational Behavior and Human Decision Processes*, September 1995, *63* (3), 287–297.

Machina, Mark J, "" Expected Utility" Analysis without the Independence Axiom," *Econometrica: Journal of the Econometric Society*, 1982, pp. 277–323.

O'Donoghue, Ted and Matthew Rabin, "Doing it now or later," *American Economic Review*, 1999, pp. 103–124.

Prelec, Drazen and George Loewenstein, "Decision Making over Time and under Uncertainty: A Common Approach," *Management Science*, July 1991, *37* (7), 770–786.

Rachlin, Howard, A.W. Logue, John Gibbon, and Marvin Frankel, "Cognition and Behavior in Studies of Choice," *Psychological Review*, 1986, *93* (1), 33–45.

_ **, Jay Brown, and David Cross**, "Discounting in Judgements of Delay and Probability," *Journal of Behavioral Decision Making*, 2000, *13* (2), 145–159.

Saito, Kota, "Strotz Meets Allais: Diminishing Impatience and the Certainty Effect: Comment," *American Economic Review*, 2011, *101* (5), 2271–75.

Segal, Uzi, "The Ellsberg Paradox and Risk Aversion: An Anticipated Utility Approach," *International Economic Review*, February 1987, *28* (1), 175–202.

Sozou, Peter D, "On hyperbolic discounting and uncertain hazard rates," *Proceedings of the Royal Society of London B: Biological Sciences*, 1998, *265* (1409), 2015–2020.

Weber, Bethany J. and Gretchen B. Chapman, "The combined effects of risk and time on choice: Does uncertainty eliminate the immediacy effect? Does delay eliminate the certainty effect?," *Organizational Behavior and Human Decision Processes*, 2005, *96*, 104–118.