AXIOMATIZATION AND MEASUREMENT OF QUASI-HYPERBOLIC DISCOUNTING*

JOSÉ LUIS MONTIEL OLEA AND TOMASZ STRZALECKI

This article provides an axiomatic characterization of quasi-hyperbolic discounting and a more general class of semi-hyperbolic preferences. We impose consistency restrictions directly on the intertemporal trade-offs by relying on what we call "annuity compensations." Our axiomatization leads naturally to an experimental design that disentangles discounting from the elasticity of intertemporal substitution. In a pilot experiment we use the partial identification approach to estimate bounds for the distributions of discount factors in the subject pool. Consistent with previous studies, we find evidence for both present and future bias. *JEL* Codes: C10, C99, D03, D90.

I. INTRODUCTION

Understanding how agents trade off costs and benefits that occur at different periods of time is a fundamental issue in economics. The leading paradigm used for the analysis of intertemporal choice has been the exponential (or geometric) discounting model introduced by Samuelson (1937) and characterized axiomatically by Koopmans (1960).

The two main properties of this utility representation are time separability and stationarity. Time separability means that the marginal rate of substitution between any two time periods is independent of the consumption levels in other periods, which rules out habit formation and related phenomena. Stationarity means that the marginal rate of substitution between any two consecutive periods is the same.

The present bias is a well-documented failure of stationarity where the marginal rate of substitution between consumption in periods 0 and 1 is smaller than the marginal rate of substitution

between periods 1 and 2. For example, the following preference pattern is indicative of present bias.

(1a) $(1, 0, 0, 0, \ldots) \succ (0, 2, 0, 0, \ldots)$

and

(1b) $(0, 1, 0, 0, 0, \ldots) \prec (0, 0, 2, 0, 0, \ldots),$

where symbols $\succ$ and $\prec$ refer to the preference over consumption streams expressed at the beginning of time before receiving any payoffs.

This article is concerned with a very widely applied model of present bias, the quasi-hyperbolic discounting model, which was first applied to study individual choice by Laibson (1997).[1] A consumption stream $(x_0, x_1, x_2, \ldots)$ is evaluated by

$$V(x_0,\, x_1,\, x_2, \ldots) = u(x_0) + \beta\delta \sum_{t=1}^{\infty} \delta^{t-1} u(x_t),$$

where u is the flow utility function, $\delta \in (0, 1)$ is the long-run discount factor, and $\beta \in (0, 1]$ is the short-run discount factor that captures the strength of the present bias; $\beta = 1$ corresponds to the standard discounted utility model.

Quasi-hyperbolic discounting retains the property of time separability but relaxes stationarity. However, the departure from stationarity is minimal: stationarity is satisfied from period $t = 1$ onward; this property is referred to as quasi-stationarity. Further relaxations of stationarity have been proposed, for instance, the generalized hyperbolic discounting of Loewenstein and Prelec (1992).[2] Our approach is not directly applicable to those models; however, both our axiomatization and experimental design extend to a class of semi-hyperbolic preferences, which approximates any time-separable preference.

Present bias may lead to violations of dynamic consistency when choices at later points in time are also part of the analysis;

---

1. This formalism was originally proposed by Phelps and Pollak (1968) to study intergenerational discounting. See also Zeckhauser and Fels (1968), published as Fels and Zeckhauser (2008).

2. Experimental studies (see, e.g., Abdellaoui, Attema, and Bleichrodt 2010; Van der Pol and Cairns 2011) find that generalized hyperbolic discounting fits the data better than the quasi-hyperbolic model. However, quasi-hyperbolic discounting is being used in many economic models, as quasi-stationarity greatly simplifies the analysis.

this has led to many different ways of modeling dynamic choice.[3] Because our results uncover the shape of "time 0" preferences without taking a stance on how they change, they can inform any of these models.

I.A. *Axiomatic Characterization*

The customary method of measuring the strength of the present bias focuses directly on the trade-off between consumption levels in periods 0 and 1; see, for example, Thaler (1981). The value of β can be revealed by varying consumption in period 1 to obtain indifference to a fixed level of consumption in time 0. However, this inference relies on parametric assumptions about the utility function u and is subject to many experimental confounds; see, for example, McClure et al. (2007), Chabris, Laibson, and Schuldt (2008), and Noor (2009, 2011) among others. Hayashi (2003) and Andersen et al. (2008) employ a conceptually related method that uses probability mixtures to elicit the trade-offs. However, this method relies on the expected utility assumption and in addition the assumption that risk aversion is inversely proportional to the elasticity of intertemporal substitution (EIS). The method that our axiomatization is building on uses only two fixed consumption levels, but instead varies the time horizon.[4] In the quasi-hyperbolic discounting model the subjective distance between periods 0 and 1 (measured by $\beta\delta$) is larger than the subjective distance between periods 1 and 2 (measured by δ), which is the reason behind the preference pattern (1a)–(1b). We uncover the parameter β by increasing the second distance enough to make it subjectively equal to the former. The delay needed to exactly match the two distances is directly related to the value of β. For example, if $\beta = \delta$, then the gap between periods 0 and 1 ($\beta\delta$) is equal to the gap between periods 1 and 3 (δ^2). In this case, the following preference pattern obtains:

$$(2a) \qquad (x, y, 0, 0, \ldots) \succ (z, w, 0, 0, \ldots)$$

---

3. For example, sophistication and naiveté (Strotz 1955), partial sophistication (O'Donoghue and Rabin 2001), costly self-control and dual-self models (Thaler and Shefrin 1981; Gul and Pesendorfer 2001, 2004; Fudenberg and Levine 2006).

4. A related but distinct method of standard sequences that relies on continuous time was used by Loewenstein and Prelec (1992) and Attema et al. (2010).

if and only if

(2b) $(0, x, 0, y, 0, 0, \ldots) \succ (0, z, 0, w, 0, 0, \ldots).$

Since we are working in discrete time, for certain values of β there may not exist a corresponding delay that would provide an exact compensation. However, it is always possible to compensate the agent with an *annuity* instead of a single payoff. For example, consider the case of $\beta = \delta + \delta^2$. In this case the simple 2-annuity provides an exact compensation:

$$(x, y, 0, 0, \ldots) \succ (z, w, 0, 0, \ldots)$$

if and only if

$$(0, x, 0, y, y, 0, \ldots) \succ (0, z, 0, w, w, 0, \ldots).$$

We show that for any β there exists an annuity that provides an exact compensation.

I.B.   *Experimental Design*

Our idea of using annuity compensations to measure impatience leads naturally to a new experimental design. Though in many cases the annuity needed for exact compensation will be very complicated, we do not insist on point-identifying the value of β. Instead, we take a simple annuity and delay it appropriately until the agent switches from "patient" to "impatient" choice. For instance, consider the following switch.

(3a) $(1, 0, 2, 2, 0, 0, 0, \ldots) \succ (2, 0, 1, 1, 0, 0, 0, \ldots)$

and

(3b) $(1, 0, 0, 2, 2, 0, 0, \ldots) \prec (2, 0, 0, 1, 1, 0, 0, \ldots).$

In comparison (3a), the agent makes the patient choice because the annuity compensation (receiving the payoff twice in a row) comes relatively soon. On the other hand, in comparison (3b), the agent makes the impatient choice because the annuity compensation is delayed. The more patient the agent, the later she switches from patient to impatient choice.

We use a multiple price list (MPL) in which we vary the delay of the annuity. The switch point from early to late rewards yields two-sided bounds on β as a function of δ. We then use the same method to elicit the value of δ. The width of the bounds on these parameters can be controlled by the length of the annuity. In our pilot study we

used the simplest 2-annuity. In Section III.B we derive two-sided bounds on the discount factors δ and β given the agent's switch point. In that section and in Appendix II we show how to use the individual bounds to partially identify the distribution of δ and β in the population. Our results are consistent with the recent experimental studies on discounting, though we treat our pilot with some caution given its online nature and lack of incentives. The partial identification methodology we develop may be useful to experimentalists using the MPL paradigm, independently of the particular preference parameters being studied.

The key aspect of our measurement method is that it disentangles discounting (as measured by β and δ) from the EIS (as measured by u). This is because we are varying the time horizon of rewards instead of varying the rewards themselves (we only use two fixed nonzero rewards). Thus, for any given β the switch point is independent of the utility function u. This is important on conceptual grounds, as impatience and EIS are separate preference parameters. By disentangling these distinct aspects of preferences, we provide a direct measurement method focusing purely on impatience.[5]

This facilitates comparisons across different rewards. It may also be useful in light of a recent debate about fungibility of rewards (see, e.g., Chabris, Laibson, and Schuldt 2008; Andreoni and Sprenger 2012; Augenblick, Niederle, and Sprenger 2013). It is often argued that observing choices over monetary payoffs is not helpful in uncovering the true underlying preferences, as those are defined on consumption, not money. Because money can be borrowed and saved, observing choices over payoff streams is informative about the shape of subjects' budget sets, but not the shape of their indifference curves. Thus, we should expect different patterns of choice between monetary and primary rewards. Because our method makes such comparisons easier, we hope that it can be used to shed some light on this issue.

The rest of the article is organized as follows. Section II presents the axioms and the representation theorems. Section III presents our method of experimental parameter measurement,

---

5. Recent experimental work has used risk preferences as a proxy for the elasticity of intertemporal substitution. However, even though these two parameters are tied together in the standard model of discounted expected utility, they are conceptually distinct (see, e.g., Epstein and Zin 1989), and there are reasons to believe they are empirically different, so one may not be a good proxy for the other.

as well as the results of a pilot study. Section IV extends our results to semi-hyperbolic discounting. Appendix I contains proofs and additional theoretical results. Appendix II contains the details of our partial identification approach. Appendix III contains additional analyses of the data and robustness checks.

II. Axiomatic Characterization

II.A. *Preliminaries*

Let C be the set of possible consumption levels, formally a connected and separable topological space. The set C could be monetary payoffs, but also any other divisible good, such as juice (McClure et al. 2007), effort (Augenblick, Niederle, and Sprenger 2013), or level of noise (Casari and Dragone 2010). Let $\mathcal{T} := \{0, 1, 2, \ldots\}$ be the set of time periods. *Consumption streams* are members of $C^{\mathcal{T}}$. A consumption stream x is *constant* if $x = (c, c, \ldots)$ for some $c \in C$. For any $c \in C$ we slightly abuse the notation by denoting the corresponding constant stream by c as well. For any $a, b, c \in C$ and $x \in C^{\mathcal{T}}$ the streams ax, abx, and $abcx$ denote $(a, x_0, x_1, \ldots), (a, b, x_0, x_1, \ldots)$, and $(a,\ b,\ c,\ x_0,\ x_1, \ldots)$, respectively.

For any T and x, y define $x_T y = (x_0, x_1, \ldots, x_T, y_{T+1}, y_{T+2}, \ldots)$. A consumption stream x is *ultimately constant* if $x = x_T c$ for some T and $c \in C$. For any T let X_T denote the set of ultimately constant streams of length T. Any X_T is homeomorphic to C^{T+1}. Consider a preference $\succsim$ defined on a subset $\mathcal{F}$ of $C^{\mathcal{T}}$ that contains all ultimately constant streams. This preference represents the choices that the decision maker makes at the beginning of time before any payoffs are realized. We focus on preferences that have a quasi-hyperbolic discounting representation over the set of streams with finite discounted utility.

Definition. A preference $\succsim$ on $\mathcal{F}$ has a *quasi-hyperbolic discounting* representation if and only if there exists a nonconstant and continuous function $u : C \to \mathbb{R}$ and parameters $\beta \in (0, 1]$ and $\delta \in (0, 1)$ such that $\succsim$ is represented by the mapping:

$$x \mapsto u(x_0) + \beta \sum_{t=1}^{\infty} \delta^t u(x_t).$$

As mentioned before, the parameter β can be thought of as a measure of the present bias. The parameter β represents the size

of the subjective distance between periods 0 and 1. As we will see, this parameter has a clear behavioral interpretation in our axiom system, and it will become explicit in what sense β is capturing the subjective distance between periods 0 and 1.

II.B. Axioms

Our axiomatic characterization involves two steps. First, by modifying the classic axiomatizations of the discounted utility model, we obtain a representation of the form:

$$(4) \qquad x \mapsto u(x_0) + \sum_{t=1}^{\infty} \delta^t v(x_t)$$

for some nonconstant and continuous $u, v : X \to \mathbb{R}$ and $0 < \delta < 1$. Second, we impose our main axiom to conclude that $v(c) = \beta u(c)$ for some $\beta \in (0, 1]$.

Our axiomatization of the representation (4) builds on the classic work of Koopmans (1960, 1972), recently extended by Bleichrodt, Rohde, and Wakker (2008). The first axiom is standard.

AXIOM 1 (WEAK ORDER). $\succsim$ is complete and transitive.

The second axiom, sensitivity, guarantees that preferences are sensitive to payoffs in periods $t = 0$ and $t = 1$ (sensitivity to payoffs in subsequent periods follows from the quasi-stationarity axiom, to be discussed later). Sensitivity is a very natural requirement, to be expected of any class of preferences in the environment we are studying.

AXIOM 2 (SENSITIVITY). There exist $e, c, c' \in C$ and $x \in \mathcal{F}$ such that $cx \succ c'x$ and $ecx \succ ec'x$.

The third axiom, initial separability, involves conditions that ensure the separability of preferences across time. (These conditions are imposed only on the few initial time periods, but extend beyond them as a consequence of the quasi-stationarity axiom.) Time separability is a necessary consequence of any additive representation of preferences and is not specific to quasi-hyperbolic discounting.

AXIOM 3 (INITIAL SEPARABILITY). For all $a, b, c, d, e, e' \in C$ and all $z, z' \in \mathcal{F}$ we have

(i) $abz \succ cdz$ if and only if $abz' \succ cdz'$,

(ii) $eabz \succ ecdz$ if and only if $eabz' \succ ecdz'$,

(iii) $ex \succ ey$ if and only if $e'x \succ e'y$.

The standard geometric discounting preferences satisfy a requirement of *stationarity*, which says that the trade-offs made at different points in time are resolved in the same way. Formally, stationarity means that $cx \succ cy$ if and only if $x \succ y$ for any consumption level $c \in C$ and streams $x, y \in \mathcal{F}$. However, as discussed in Section I, the requirement of stationarity is not satisfied by quasi-hyperbolic discounting preferences; in fact, it is the violation of stationarity that is often taken to be synonymous with quasi-hyperbolic discounting. Nevertheless, quasi-hyperbolic discounting preferences possess strong stationarity-like properties, since the preferences starting from period 1 onward are geometric discounting.

AXIOM 4 (QUASI-STATIONARITY). For all $e, c \in C$ and all $x, y \in \mathcal{F}$, $ecx \succ ecy$ if and only if $ex \succ ey$.

The last three axioms, introduced by Bleichrodt, Rohde, and Wakker (2008), are used instead of stronger infinite-dimensional continuity requirements. They are of technical nature, as are all continuity-like requirements. However, constant equivalence and tail continuity have simple interpretations in terms of choice behavior.

AXIOM 5 (CONSTANT-EQUIVALENCE). For all $x \in \mathcal{F}$ there exists $c \in C$ such that $x \sim c$.

AXIOM 6 (FINITE CONTINUITY). For any T, the restriction of $\succ$ to X_T satisfies continuity, that is, for any $x \in X_T$ the sets $\{y \in X_T : y \succ x\}$ and $\{y \in X_T : y \prec x\}$ are open.

AXIOM 7 (TAIL-CONTINUITY). For any $c \in C$ and any $x \in \mathcal{F}$ if $x \succ c$, then there exists τ such that for all $T \geq \tau$, $x_T c \succ c$; if $x \prec c$, then there exists τ such that for all $T \geq \tau$, $x_T c \prec c$.

THEOREM 1. The preference $\succsim$ satisfies Axioms 1–7 if and only if it is represented by equation (4) for some nonconstant and continuous $u, v : X \to \mathbb{R}$ and $0 < \delta < 1$.

Note that the representation obtained in Theorem 1 is a generalization of the quasi-hyperbolic model. The main two features of this representation are the intertemporal separability of consumption and the standard stationary behavior that follows period 1 (captured by the quasi-stationarity axiom).

The restriction that specifies representation (4) to the quasi-hyperbolic class imposes a strong relationship between the utility functions u and v. Not only do they have to represent the same ordering over the consumption space C, they also must preserve the same cardinal ranking, that is, u and v relate to each other through a positive affine transformation $u = \beta v$ (the additive constant can be omitted without loss of generality). To capture this restriction behaviorally, we express it in terms of the willingness to make trade-offs between time periods.

We now present three different ways of restricting equation (4) to the quasi-hyperbolic model. It is important to observe that an axiom that requires the preference relation $\succsim$ to exhibit preference pattern (1) is necessary but not sufficient to pin down the $\beta\delta$ model: present bias may arise as an immediate consequence of different preference intensity—as captured by differences in u and v. Therefore, in the context of representation (4), present bias could be explained without relying on the $\beta\delta$ structure. The additional axioms that we propose shed light on what it exactly means, in terms of consumption behavior, to have different short-term discount factors and a common utility index.

II.C. *The Annuity Compensation Axiom*

First, we present an axiom that ensures δ is larger than half. We impose this requirement to be able to construct a "future compensation scheme" that exactly offsets the lengthening of the first time period caused by β. If δ is less than half, then there will be values of β which we cannot compensate for exactly.[6]

AXIOM 8 ($\delta \geq 0.5$). If $(c, a, a, \ldots) \succ (c, b, b, \ldots)$ for some $a, b, c \in X$, then

$$(c, b, a, a, \ldots) \succsim (c, a, b, b \ldots).$$

In the context of representation (4) the long-run patience (δ) can be easily measured. Fix two elements $a, b \in C$ such that a is preferred to b. Axiom 8 uncovers the strength of patience by getting information about the following trade-off. Consider first a consumption stream that pays a tomorrow and b forever after. Consider now a second consumption stream in which the order of the alternatives is reversed. An agent who decides to postpone

---

6. Since in most calibrations δ is close to 1 for any reasonable length of the time period, we view this step as innocuous.

higher utility (by choosing b first) reveals a certain degree of patience. Under representation (4) the patient choice reveals a value of $\delta \geq 0.5$.

THEOREM 2. Suppose $\succsim$ is as in Theorem 1. It satisfies Axiom 8 if and only $\delta \geq 0.5$.

As discussed in Section I, our main axiom relies on the idea of increasing the distance between future payoffs to compensate for the lengthening of the time horizon caused by β. For example, if $\beta = \delta$, then the trade-off between periods 0 and 1 is the same as the trade-off between periods 1 and 3. Similarly, if $\beta = \delta^t$, then the trade-off between periods 0 and 1 is the same as the trade-off between periods 1 and $t + 2$. Because we are working in discrete time, there exist values of β such that $\delta^{t+1} < \beta < \delta^t$ for some t, so that the exact compensation by one payoff is not possible. However, due to time separability, it is possible to compensate the agent by an annuity. Lemma 1 in the Appendix shows that as long as $\delta \geq 0.5$, any value of β can be represented by a sum of the powers of δ with coefficients 0 or 1.[7] The set M is the collection of powers with nonzero coefficients; formally, let M denote a subset of $\{2, 3, \ldots\} \subseteq \mathcal{T}$. We refer to M as an annuity. Our main axiom guarantees that the annuity M is independent of the consumption levels used to elicit the trade-offs.

AXIOM 9 (ANNUITY COMPENSATION). There exists an annuity M such that for all a, b, c, d, e

$$\begin{pmatrix} a & \text{if } t = 0 \\ b & \text{if } t = 1 \\ e & \text{otherwise} \end{pmatrix} \succ \begin{pmatrix} c & \text{if } t = 0 \\ d & \text{if } t = 1 \\ e & \text{otherwise} \end{pmatrix}$$

if and only if

$$\begin{pmatrix} a & \text{if } t = 1 \\ b & \text{if } t \in M \\ e & \text{otherwise} \end{pmatrix} \succ \begin{pmatrix} c & \text{if } t = 1 \\ d & \text{if } t \in M \\ e & \text{otherwise} \end{pmatrix}.$$

The main result of our article is the following theorem.

7. A similar technique was used in repeated games, see, for example, Sorin (1986) and Fudenberg and Maskin (1991). We thank Drew Fudenberg for these references. See also Kochov (2013), who uses results from number theory to calibrate the discount factor in the geometric discounting model.

THEOREM 3. A preference $\succsim$ satisfies Axioms 1–9 if and only if has a quasi-hyperbolic discounting representation with $\delta \geq 0.5$. In this case, $\beta = \sum_{t \in M} \delta^{t-2}$.

II.D. Alternative Axioms

The annuity compensation axiom ensures that v is cardinally equivalent to u. From the formal logic viewpoint, however, the compensation axiom involves an existential quantifier. This section complements our analysis by considering two alternate ways of ensuring the cardinal equivalence: a form of the trade-off consistency axiom and a form of the independence axiom.

Both axioms need to be complemented with an axiom that guarantees that $\beta < 1$. The following axiom yields just that.

AXIOM 10 (PRESENT BIAS). For any $a, b, c, d, e \in C$, $a \succ c$

$$(e, a, b, e \dots) \sim (e, c, d, e, \dots) \Longrightarrow (a, b, e, \dots) \succsim (c, d, e, \dots).$$

This axiom says that if two distant consumption streams are indifferent, one "impatient" (involving a bigger prize at $t = 1$, followed by a smaller at $t = 2$) and one "patient" (involving a smaller prize at $t = 1$, followed by a bigger at $t = 2$), then pushing both of them forward will skew the preference toward the impatient choice.

For both approaches, fix a consumption level $e \in C$ (for example in the context of monetary prizes, e could be \$0). For any pair of consumption levels $a, b \in C$ let (a, b) denote the consumption stream $(a, b, b, b, \dots)$.

1. Trade-off Consistency Axiom.

AXIOM 11 (TRADE-OFF CONSISTENCY). For any $a, b, c, d, e_1, e_2 \in C$,

If $(b, e_2) \succsim (a, e_1)$, $(c, e_1) \succsim (d, e_2)$, and $(e_3, a) \sim (e_4, b)$, then $(e_3, c) \succsim (e_4, d)$,

and

If $(e_2, b) \succsim (e_1, a)$, $(e_1, c) \succsim (e_2, d)$, and $(a, e_3) \sim (b, e_4)$, then $(c, e_3) \succsim (d, e_4)$.

The intuition behind the first requirement of axiom is as follows (the second requirement is analogous and ensures that the time periods are being treated symmetrically). The first premise is that the "utility difference" between b and a offsets the utility difference between e_1 and e_2. The second premise is that the

utility difference between e_1 and e_2 offsets the utility difference between d and c. These two taken together imply that the utility difference between b and a is bigger than the utility difference between d and c. Thus, if the utility difference between e_3 and e_4 exactly offsets the utility difference between b and a, it must be big enough to offset the utility difference between d and c.

THEOREM 4. The preference $\succsim$ satisfies Axioms 1–7 and 11 if and only if there exists a nonconstant and continuous function $u : C \to \mathbb{R}$ and parameters $\beta > 0$ and $\delta \in (0, 1)$ such that $\succsim$ is represented by the mapping:

$$x \mapsto u(x_0) + \beta \sum_{t=1}^{\infty} \delta^t u(x_t).$$

Moreover, it satisfies Axiom 10 if and only if $\beta \le 1$, that is, $\succsim$ has the quasi-hyperbolic discounting representation.

2. Independence Axiom. By continuity (Axioms 6 and 7) for any $a, b \in C$ there exists a consumption level c that satisfies $(c, c) \sim (a, b)$. Let $c(a, b)$ denote the set of such consumption levels. Note that we are not imposing any monotonicity assumptions on preferences (the set C could be multidimensional), and for this reason the set $c(a, b)$ may not be a singleton. However, since all of its members are indifferent to each other, it is safe to assume in the expressions below that $c(a, b)$ is an arbitrarily chosen element of that set.

AXIOM 12 (INDEPENDENCE). For any $a, a', a'', b, b', b'' \in C$ if $(a, b) \succsim (a', b')$, then

$$(c(a, a''), c(b, b'')) \succsim (c(a', a''), c(b', b''))$$

and

$$(c(a'', a), c(b'', b)) \succsim (c(a'', a'), c(b'', b')).$$

The intuition behind the first requirement of the axiom is as follows (the second requirement is analogous and ensures that the time periods are being treated symmetrically). For any $(a, b), (a'', b'')$ the stream given by $(c(a, a''), c(b, b''))$ is a "subjective mixture" of bets (a, b) and (a'', b''). The axiom requires that if one

consumption stream is preferred to another, then mixing each stream with a third stream preserves the preference.[8]

The next axiom, is a version of Savage's P3. It ensures that preferences in each time period are ordinally the same.

AXIOM 13 (MONOTONICITY). *For any* $a, b, e \in C$, *then*

$$b \succsim a \Longleftrightarrow (b, e) \succsim (a, e) \text{ and } (e, b) \succsim (e, a).$$

THEOREM 5. *The preference* $\succsim$ *satisfies Axioms 1–7 and 12–13 if and only if there exists a nonconstant and continuous function* $u : C \to \mathbb{R}$ *and parameters* $\beta > 0$ *and* $\delta \in (0, 1)$ *such that* $\succsim$ *is represented by the mapping:*

$$x \mapsto u(x_0) + \beta \sum_{t=1}^{\infty} \delta^t u(x_t).$$

Moreover, it satisfies Axiom 10 if and only if $\beta \leq 1$, *that is,* $\succsim$ *has the quasi-hyperbolic discounting representation.*

II.E. *Related Theoretical Literature*

A large part of the theoretical literature on time preferences uses the choice domain of dated rewards, where preferences are defined on $C \times \mathcal{T}$, that is, only one payoff is made. On this domain Fishburn and Rubinstein (1982) axiomatized exponential discounting. By assuming that $\mathcal{T} = \mathbb{R}_+$, that is, that time is continuous, Loewenstein and Prelec (1992) axiomatized a generalized model of hyperbolic discounting, where preferences are represented by $V(x, t) = (1 + \alpha t)^{-\frac{\beta}{\alpha}} u(x)$. Recently, Attema et al. (2010) generalized this method and obtained an axiomatization of quasi-hyperbolic discounting, among other models.

The foregoing results share a common problem: the domain of dated rewards is not rich enough to enable the measurement of the levels of discount factors. Even in the exponential discounting model the value of δ can be chosen arbitrarily, as long as it belongs to the interval $(0, 1)$, see, for example, Theorem 2 of Fishburn and Rubinstein (1982); see also the recent results of Noor (2011).

---

8. We thank Simon Grant for suggesting this type of axiom. A similar approach along the lines of Nakamura (1990) is considered in the Appendix.

The richer domain of consumption streams that we employ in this article allows us to elicit more complex trade-offs between time periods and to pin down the value of all discount factors.

The continuous time approach can be problematic for yet another reason. It relies on extracting a sequence of time periods of equal subjective length, a so-called standard sequence.[9] Since the time intervals in a standard sequence are of equal subjective length, their objective duration is unequal and has to be uncovered by eliciting indifferences. In contrast, our method uses time intervals of objectively equal length and does not rely on such elicitation.

Finally, an axiomatization of quasi-hyperbolic discounting using a different approach was obtained by Hayashi (2003). He studied preferences over an extended domain that includes lotteries over consumption streams. He used the lottery mixtures to calibrate the value of β. His axiomatization and measurement rely heavily on the assumption of expected utility, which is rejected by the bulk of experimental evidence. Moreover, in his model the same utility function u measures both risk aversion and the intertemporal elasticity of substitution; however, these two features of preferences are conceptually unrelated (see, e.g., Kreps and Porteus 1978; Epstein and Zin 1989) and are shown to be different in empirical calibrations. Another limitation of his paper is that his axioms are not suggestive of a measurement method of the relation between the short-run and long-discount factor.

III. EXPERIMENTAL DESIGN AND A PILOT STUDY

In this section we use the idea of annuity compensations that underlies our axiomatization and provide a preference elicitation design. The method provides two-sided bounds for β_i and δ_i for each subject i. Because there is a natural heterogeneity of preferences in the population, we are interested not only in average values, but in the whole distribution. We use these bounds to partially identify the cumulative distribution functions of β_i and

---

9. The standard sequence method was originally applied to eliciting subjective beliefs by Ramsey (1931) and later by Luce and Tukey (1964). Interestingly, the similarity between beliefs and discounting was already anticipated by Ramsey: "the degree of belief is like a time interval; it has no precise meaning unless we specify how it is to be measured."

δ_i in the population. Our method works independently of the utility function, so no functional form assumptions have to be made and no curvatures have to be estimated. We first discuss the design and then report results of a pilot experiment.

III.A. Design

The proposed experiment provides a direct test of stationarity; moreover, under the assumption that agent i's preferences belong to the quasi-hyperbolic class, our experimental design yields two-sided bounds on the discount factors β_i and δ_i.[10] The size of the bounds depends on the choice of the annuity M. We use the simplest annuity composed of just two consecutive payoffs; however, tighter measurements are possible. The individual bounds are used to partially identify the (marginal) distributions of preference parameters δ_i and β_i in the population. All the details concerning the partial identification of the marginal distributions are provided in Appendix II.

As mentioned before, the experiment does not rely on any assumptions about the curvature of the utility function u_i. In fact, whether the prizes are monetary is immaterial; the only assumption that the researcher has to make is that there exist two prizes a and b, where b is more preferred than a (it doesn't matter by how much). As a consequence, the experimental design can be used to study how the nature of the prize (e.g., money, effort, consumption good, addictive good) affects impatience, a feature not shared by experiments based on varying the amount of monetary payoff.

The questionnaire consists of two MPLs.[11] In each list, every question is a choice between two consumption plans: A (impatient choice) and B (patient choice); see, for example, Figures I and II. Each option in the first list involves an immediate payoff followed by a two-period annuity that pays off the same outcome in periods t and $t+1$; the second list is a repetition of the first with all payoffs delayed by one period. Under the assumption of quasi-hyperbolic discounting, the agent has only one switch point in each list, that

10. In principle, all our axioms are testable, so that assumption could be verified as well.

11. MPLs have been used to elicit discount factors for some time now. For example, Coller and Williams (1999) and Harrison, Lau, and Williams (2002) use them under the assumption of linear utility and geometric discounting. Andreoni, Kuhn, and Sprenger (2013) use them under the assumption of CRRA utility.
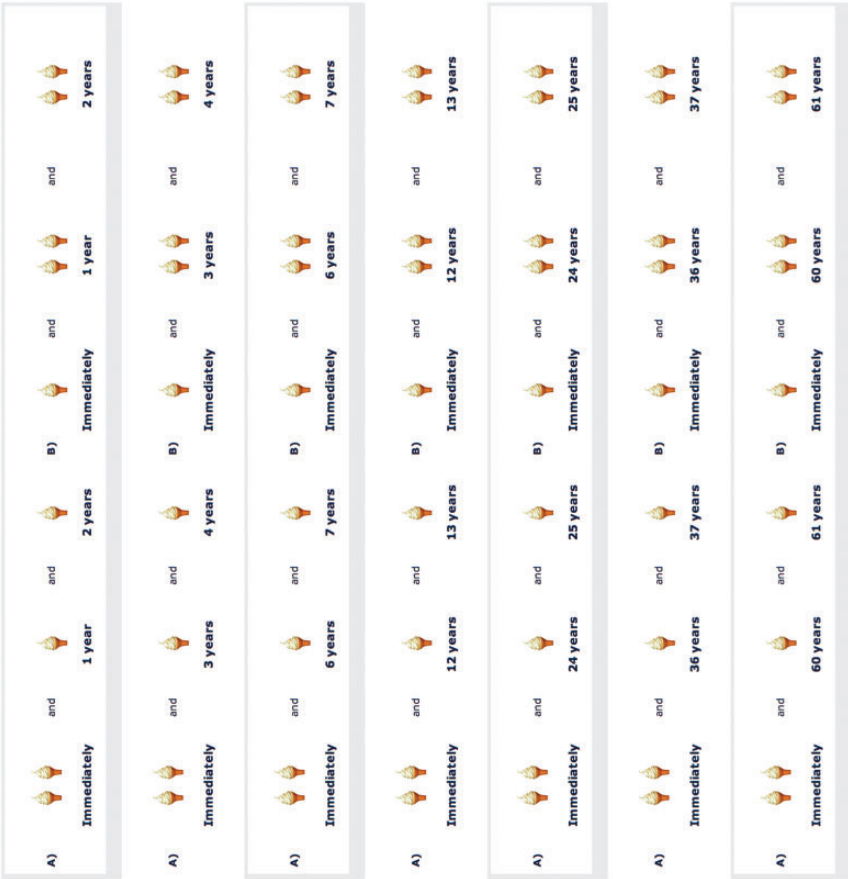
FIGURE I

First Price List

FIGURE II

Second Price List

is, she answers B for questions $1, \ldots, k$ and A for questions $k+1, \ldots, n$ (where n is the total number of questions in the list).[12]

III.B. *Parameter Bounds*

Since the second list does not involve immediate payoffs, the observed switch point in this list (denoted $s_{i,2}$) yields bounds on the discount factor δ_i. For example, suppose that in the list depicted in Figure II subject i chose B in the first five questions and A in all subsequent questions, so that $s_{i,2} = 6$. Then,

$$\beta_i \delta_i u_i(1) + \beta_i \delta_i^{25} u_i(2) + \beta_i \delta_i^{26} u_i(2) \geq \beta_i \delta_i u_i(2) + \beta_i \delta_i^{25} u_i(1) + \beta_i \delta_i^{26} u_i(1)$$
$$\beta_i \delta_i u_i(1) + \beta_i \delta_i^{37} u_i(2) + \beta_i \delta_i^{38} u_i(2) \leq \beta_i \delta_i u_i(2) + \beta_i \delta_i^{37} u_i(1) + \beta_i \delta_i^{38} u_i(1),$$

where $u(2)$ is the utility of two ice cream cones and $u(1)$ is the utility of one cone. If $u(2) > u(1)$ this is equivalent to $\delta_i^{36} + \delta_i^{37} \leq 1 \leq \delta_i^{24} + \delta_i^{25}$, so approximately

$$0.972 \leq \delta_i \leq 0.981.$$

Therefore, the probability of the event $\{i \,|\, s_{i,2} = 6\}$ provides a lower bound for the probability of the event $\{i \,|\, 0.972 \leq \delta_i \leq 0.981\}$. Appendix II derives upper and lower bounds for the marginal distribution of δ_i based on the switch point $s_{i,2}$.

Note that if the switch points in the first and second list are different, stationarity is violated and we obtain bounds on β_i. For example, suppose that in the first price list the subject answered B in the first three questions and A in all subsequent questions, so that $s_{i,1} = 4$. We have

$$u_i(1) + \beta_i \delta_i^6 u_i(2) + \beta_i \delta_i^7 u_i(2) \geq u_i(2) + \beta_i \delta_i^6 u_i(1) + \beta_i \delta_i^7 u_i(1)$$
$$u_i(1) + \beta_i \delta_i^{12} u_i(2) + \beta_i \delta_i^{13} u_i(2) \leq u_i(2) + \beta_i \delta_i^{12} u_i(1) + \beta_i \delta_i^{13} u_i(1),$$

or equivalently, $s_{i,1} = 4$ implies

$$\frac{1}{\delta_i^6 + \delta_i^7} \leq \beta_i \leq \frac{1}{\delta_i^{12} + \delta_i^{13}},$$

so using the bounds for δ_i just derived from the second list we conclude that $s_{i,1} = 4$ and $s_{i,2} = 6$ imply

$$0.565 \leq \beta_i \leq 0.712.$$

---

12. In fact, the switch point is unique under any time-separable model à la Ramsey (1931) with a representation $\sum_{t=0}^{\infty} D_t u(c_t)$, where $D_{t+1} < D_t$, for example the generalized hyperbolic discounting model of Loewenstein and Prelec (1992).

Appendix II derives upper and lower bounds for the marginal distribution of β_i based on the switch points $s_{i,1}$ and $s_{i,2}$.

III.C. *Implementation of the Pilot Experiment*

To illustrate our design, we implemented a pilot study using an online platform and hypothetical rewards. Though comparative studies show that there tends to be little difference between choices with hypothetical and real consequences in discounting tasks (Johnson and Bickel 2002) and that online markets provide good quality data and replicate many lab studies (Horton, Rand, and Zeckhauser 2011), we treat our results with caution and think of this study as a proof of concept before a thorough incentivized laboratory or field experiment can be implemented.[13] We use two kinds of hypothetical rewards: money and ice cream. We have a total of 1,277 participants each with a unique IP address; 639 subjects answered the money questionnaire and 640 the ice cream questionnaire (548 participants answered both).

The experiment was conducted using Amazon's Mechanical Turk (AMT), an online labor market. Immediate and convenient access to a large and diverse subject pool is usually emphasized as one of the main advantages of the online environment; see, for example, Mason and Suri (2012). One of the common concerns often raised by online experiments is that both low wages and the lack of face-to-face detailed instructions to participants might lead to low-quality answers. However, Mason and Watts (2010), Mason and Suri (2012), and Marge, Banerjee, and Rudnicky (2010) present evidence of little to no effect of wage on the quality of answers, at least for some kinds of tasks. In our study we paid $5 per completed questionnaire. The average duration of each questionnaire was five minutes. Hence, we paid approximately $60 per hour: a significantly larger reward than the reservation wage of $1.38 per hour reported in Mason and Suri (2012) for AMT workers.

The lack of face-to-face detailed instructions is often addressed by creating additional questions to verify subjects' understanding of the experiment (Paolacci, Chandler, and Ipeirotis 2010). To address these concerns, we have two questions at the beginning of the questionnaire that check participants'

---

13. Hypothetical rewards may offer some benefits compared to real rewards because they eliminate the need for using front-end delays so the "present moment" in the lab is indeed present.

understanding. Out of the 638 (639) participants in the money treatment, a subsample of 502 (503) subjects was selected based on "monotonicity" and "understanding" initial checks; see the Online Appendix.

We also perform two additional robustness checks: we study response times and we vary worker qualifications. These exercises are described in the Online Appendix.

An important consideration when using the MPL paradigm are multiple switch points. As noted in Section III.A, any agent with a time-separable impatient preference has a unique switch point. Three hundred thirty-six out of the 502 subjects in the money treatment and 444 out of the 503 subjects in the ice cream treatment have a unique switch point. We focus only on those subjects, disregarding the multiple switchers.

We note that there is an important share of "never switchers" in our sample; that is, subjects who always chose the patient (or impatient) prospect in both price lists. Since never switchers are compatible with both $\beta_i \leq 1$ and $\beta_i \geq 1$, they directly affect the width of our bounds for the cumulative distribution function (c.d.f.) of β. We did not disregard never switchers, as we have no principled way of doing so: their response times were not significantly faster than those of the subjects that exhibited a switch point and the fraction of such subjects was independent of the worker qualifications (for details see the Online Appendix). In small-scale pilot tests with shorter time horizons, even more subjects were never switching, which is what prompted us to use longer time horizons.[14] We are hopeful that the number of never switchers will decrease in the lab and/or with real incentives, which would allow for more practical time horizons.[15]

III.D. *Results of the Experiment*

As discussed in Section III.B, for each such subject, we obtain two-sided bounds on δ_i; we use these bounds to partially identify the distribution of δ in the population. To represent the aggregate distribution of δ in our subject population we graph two

14. Dohmen et al.'s (2012) experiment shows that the elicited preferences can depend on the time horizon. The dependence can be so strong that it leads to intransitives.

15. However, we note that similar behavior was obtained in the lab with real incentives by Andreoni and Sprenger (2012), where in a convex time budget task roughly 70% of responses were corner solutions and 37% of subjects never chose interior solutions.

nondecreasing functions, each corresponding to one of the ends of the interval. The true c.d.f. must lie in between them. Figure III presents the c.d.f. bounds for the two treatments; the true c.d.f. must lie in the gray area between the dashed line (upper bound) and the solid line (lower bound).

We now turn to β. As discussed in Section III.B, for each subject we obtain two-sided bounds on β_i using his answers in the first price list and bounds on his δ_i obtained above. We use the same method of aggregating these bounds as above. Figure IV presents the c.d.f. bounds for the two treatments; once again, the true c.d.f. must lie in the gray area. We reiterate that obtaining tighter bounds on the distribution of β is possible by using annuity compensation schemes longer than the simple two-period annuity that we adopted here for simplicity.

The distribution of parameter values seems consistent with results in the literature. The next section makes detailed comparisons. A noticeable feature of the data is the high proportion of subjects with $\beta > 1$, that is, displaying a future bias. This has been documented by other researches as well; for example, Read (2001), Gigliotti and Sopher (2003), Scholten and Read (2006), Sayman and Öncüler (2009), Attema et al. (2010), Cohen, Tallon, and Vergnaud (2011), Takeuchi (2011), Andreoni and Sprenger (2012), and Halevy (2012).

III.E. Relation to the Experimental Literature

There is a large body of research on estimation of time preferences using laboratory experiments. The picture that seems to emerge is that little present bias is observed in studies using money as rewards, whereas it emerges strongly in studies using primary rewards. For example, Andreoni and Sprenger (2012) introduce the convex time budget procedure to jointly estimate the parameters of the β-δ model with constant relative risk aversion (CRRA) utility. They find averages values of δ between 0.74 and 0.8, and only 16.7% of their subjects exhibit diminishing impatience. The null hypothesis of exponential discounting, $\beta = 1$, is rejected against the one-sided alternative of future bias, $\beta > 1$. Andreoni, Kuhn, and Sprenger (2013) compare the convex time budget procedure and what they call dual marginal price lists in the context of the CRRA discounted utility model. Even though they find substantial difference in curvature estimates arising from the two methodologies, they find similar time preference

(a)

(b)

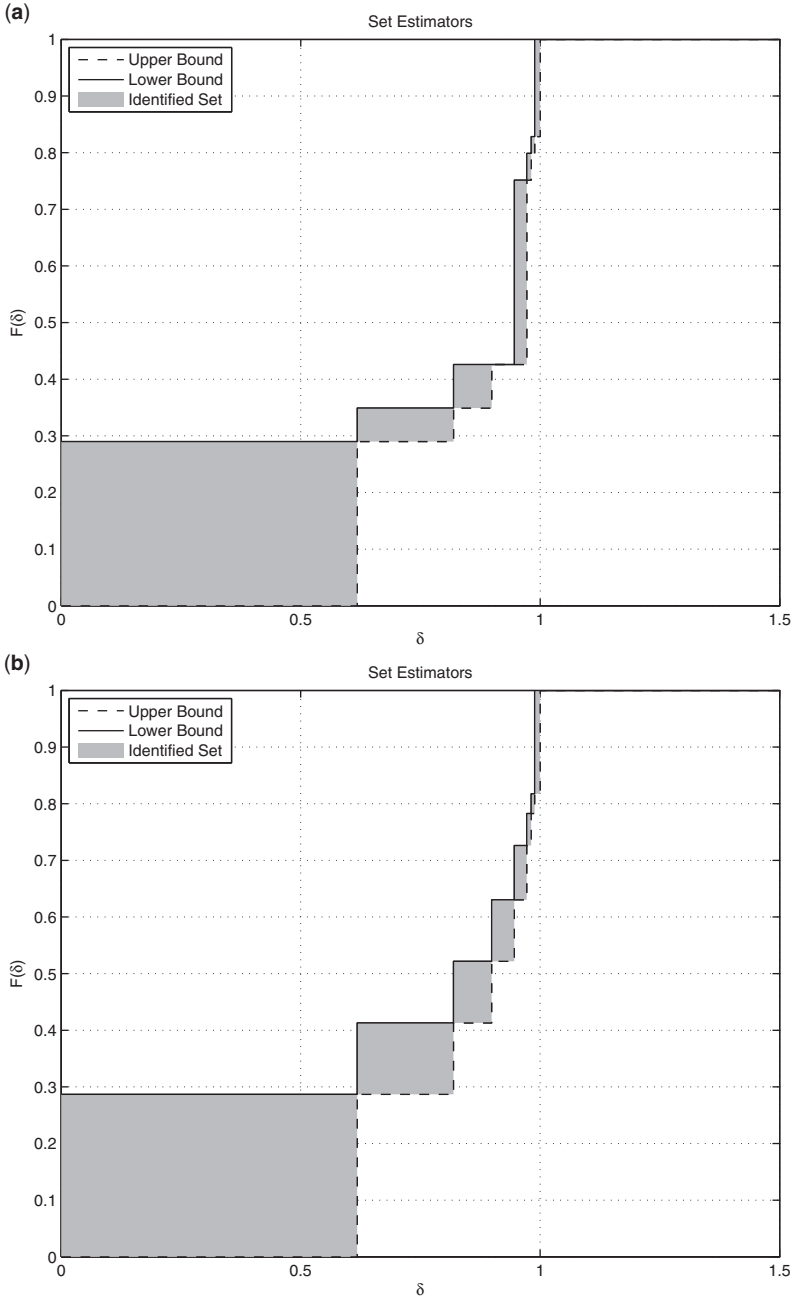
FIGURE III

Bounds for the c.d.f. of δ: a) Money b) Ice cream

(a)

Bounds for the c.d.f. of β:
Set Estimators

(b)

Bounds for the c.d.f. of β:
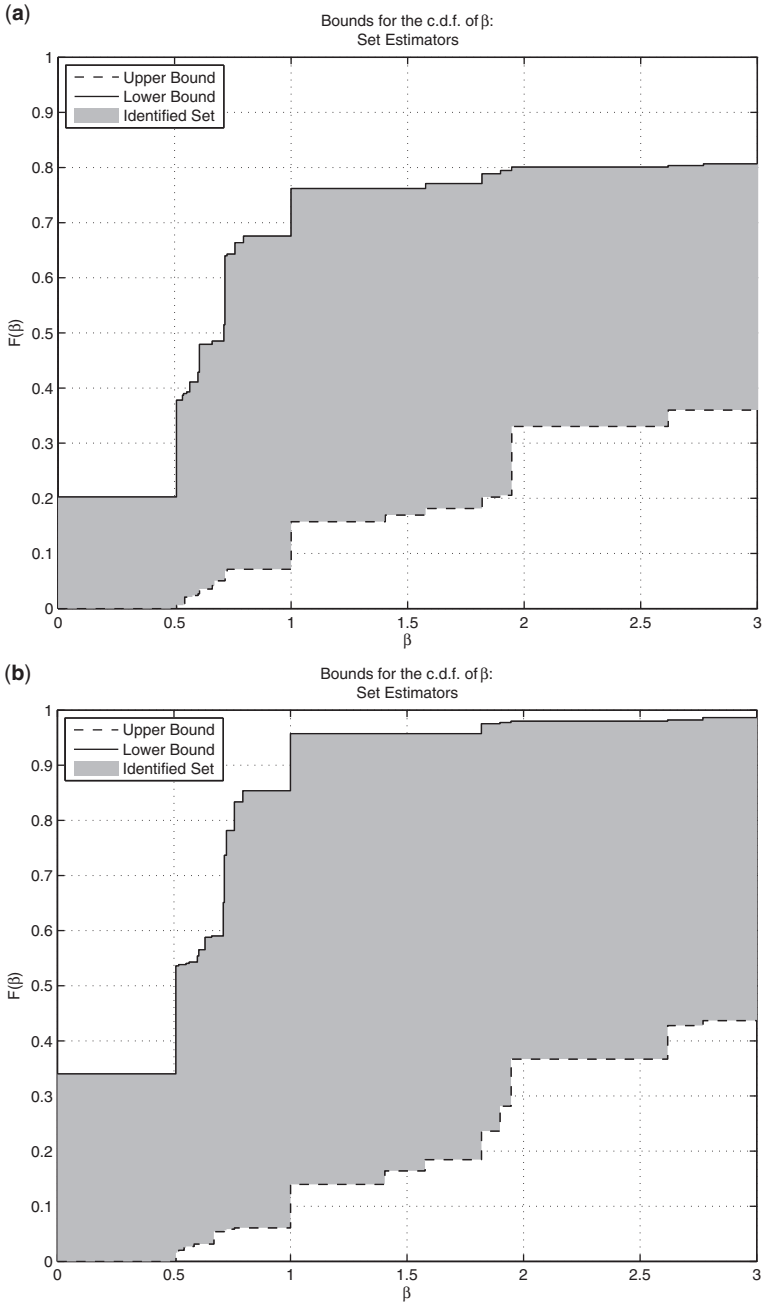Set Estimators

FIGURE IV

Bounds for the c.d.f. of β: a) Money b) Ice cream

parameters. The reported estimates of yearly δ are around 0.7. They again find very little evidence of quasi-hyperbolic discounting. Using risk aversion as proxy for the EIS Andersen et al. (2008) find that 72% of their subjects are exponential and 28% are hyperbolic.

Another line of work relies on a parameter-free measurement of utility. Using hypothetical rewards and allowing for differential discounting of gains and losses Abdellaoui, Attema, and Bleichrodt (2010) show that generalized hyperbolic discounting fits the data better than exponential discounting and quasi-hyperbolic discounting, where the median values of β are close to 1. In an innovative experiment Halevy (2012) elicits dynamic choices to study the present bias, as well as time consistency and time invariance of preferences. Since we only focus on time 0 preferences, only his results on the present bias are relevant to us. He finds that 60% of his subjects have stationary preferences, 17% display present bias, and 23% display future bias.

On the other hand, the present bias is strong in studies using primary rewards. For example, McClure et al. (2007) use fruit juice and water as rewards and find that on average $\beta \approx 0.52$. Augenblick, Niederle, and Sprenger (2013) compare preference over monetary rewards and effort. Using parametric specifications for both utility functions, they show little present bias for money, but existing present bias for effort: they find that for money the average $\beta \approx 0.98$ but ranges between 0.87 and 0.9 for effort (depending on the task). Using health outcomes as rewards, Van der Pol and Cairns (2011) find significant violations of stationarity (however, their results point in the direction of generalized hyperbolic, rather than quasi-hyperbolic discounting).

Turning to our experiment, the results of our money treatment are consistent with those already mentioned, that is, the present bias is not prevalent: at least 10% of subjects have $\beta < 1$ and at least 30% of subjects have $\beta > 1$. Our second treatment used a primary reward—ice cream—in the hope of obtaining a differential effect. However, the effect is weak: at least 10% of subjects have $\beta < 1$ and at least 10% of subjects have $\beta > 1$. This is consistent with the average β being lower for primary rewards. A possible explanation of the weakness of the effect is that hypothetical rewards may lead subjects to conceptualize money and ice cream similarly. A larger difference would more likely be seen in a study using real incentives.

IV. Semi-hyperbolic Preferences

As mentioned earlier, other models of the present bias relax stationarity beyond the first time period. The most general model that maintains time separability is one where

$$V(x_0, x_1, \ldots) = \sum_{t=0}^{\infty} D_t u(x_t),$$

where $1 = D_0 > D_1 > \cdots > 0$. For these preferences to be defined on constant consumption streams the condition $\sum_{t=0}^{\infty} D_t < \infty$ has to be satisfied. We call this class *time-separable preferences* (TSP). An example of TSP is the generalized hyperbolic discounting model of Loewenstein and Prelec (1992) where $D_t = (1 + \alpha t)^{-\frac{\beta}{\alpha}}$ and $\beta > \alpha$.

Consider the subclass of *semi-hyperbolic preferences*, where $D_1, \ldots, D_T$ are unrestricted and for some $\delta \in (0, 1)$, $\frac{D_{t+1}}{D_t} = \delta$ for all $t > T$. This class does not impose any restrictions on the discount factors for a finite time horizon and assumes that they are exponential thereafter. Notice that if the time horizon is finite, this implies that semi-hyperbolic preferences coincide with TSP. We now show that with infinite horizon semi-hyperbolic preferences approximate any TSP for bounded consumption streams. We say that a stream $x = (x_0, x_1, \ldots)$ is *bounded* whenever there exist $\underline{c}, \overline{c} \in C$ such that $\underline{c} \preccurlyeq x_t \preccurlyeq \overline{c}$ for all t. The restriction to bounded plans may be a problem in models where economic growth is unbounded, but seems realistic in experimental settings.

THEOREM 6. For any V that belongs to the TSP class there exists a sequence V^n of semi-hyperbolic preferences such that $V^n(x) \to V(x)$ for all bounded x. Moreover, the convergence is uniform on any set of equi-bounded consumption streams. Furthermore, this implies that: (i) if $x \succsim^n y$ for all n sufficiently large, then $x \succsim y$ and (ii) if $x \succ y$ then for all n large enough $x \succ^n y$.

To extend our axiomatization to semi-hyperbolic preferences, quasi-stationarity, initial separability, and annuity compensation need to be modified. Quasi-stationarity needs to be relaxed to hold starting from period T. Initial separability needs to be be imposed for periods $t = 0, 1, \ldots T$ instead of just 0, 1, 2 (this property was implied by initial separability

together with quasi-stationarity, but the latter axiom is now weaker, so it has to be assumed directly). Annuity compensation becomes:

AXIOM 14 (EXTENDED ANNUITY COMPENSATION). For each $\tau = 0$, $1, \ldots T$ there exists an annuity M such that for all a, b, c, d, e

$$\begin{pmatrix} a & \text{if } t = \tau \\ b & \text{if } t = \tau + 1 \\ e & \text{otherwise} \end{pmatrix} \succ \begin{pmatrix} c & \text{if } t = \tau \\ d & \text{if } t = \tau + 1 \\ e & \text{otherwise} \end{pmatrix}$$

if and only if

$$\begin{pmatrix} a & \text{if } t = T + 1 \\ b & \text{if } t \in M \\ e & \text{otherwise} \end{pmatrix} \succ \begin{pmatrix} c & \text{if } t = T + 1 \\ d & \text{if } t \in M \\ e & \text{otherwise} \end{pmatrix}.$$

Finally, to understand how to extend our experimental design to semi-hyperbolic preferences, consider the following generalization of quasi-hyperbolic discounting, the α-β-δ preferences, where

$$V(x_0, x_1, \ldots) = u(x_0) + \alpha\beta\delta \left[u(x_1) + \beta\delta \sum_{t=2}^{\infty} \delta^{t-2} u(x_t) \right].$$

The elicitation of δ is from an MPL like in Figure II, where the first payoff is in two years instead of one year. The elicitation of β is from an MPL like in Figure II. The elicitation of α is from an MPL like in Figure I. The practicality of this approach depends on how well the semi-hyperbolic preferences approximate the observed preferences for reasonable time horizons. This is an empirical question beyond the scope of this article.

V. CONCLUSION

This article axiomatizes the class of quasi-hyperbolic discounting and provides a measurement technique to elicit the preference parameters. Both methods extend to what we call semi-hyperbolic preferences. Both methods are applications of the same basic idea: calibrating the discount factors using annuities. In the axiomatization we are looking for an exact compensation, whereas in the experiment we use a multiple price list to get two-sided bounds. The advantage of this method is that it

disentangles discounting from the EIS and hence facilitates comparisons of impatience across rewards. To illustrate our experimental design we run an online pilot experiment using the β-δ model. We show how to partially identify the distribution of discount factors in the population.

APPENDIX I: PROOFS

Proof of Theorem 1

Necessity of the axioms is straightforward. For sufficiency, we follow a sequence of steps.

Step 1. The initial separability axiom guarantees that the sets $\{0, 1\}, \{1, 2\},$ and $\{1, 2, \ldots, \}$ are independent. To show that for all $t = 2, \ldots$ the sets $\{t, t + 1\}$ are independent fix $x, y, z, z' \in \mathcal{F}$ and suppose that

$$(z_0, z_1, \ldots, z_{t-1}, x_t, x_{t+1}, z_{t+1}, \ldots) \succ (z_0, z_1, \ldots, z_{t-1}, y_t, y_{t+1}, z_{t+1}, \ldots).$$

Apply quasi-stationarity $t - 1$ times to obtain

$$(z_0, x_t, x_{t+1}, z_{t+1}, \ldots) \succ (z_0, y_t, y_{t+1}, z_{t+1}, \ldots).$$

By part (ii) of initial separability, conclude that

$$(z_0, x_t, x_{t+1}, z'_{t+1}, \ldots) \succ (z_0, y_t, y_{t+1}, z'_{t+1}, \ldots).$$

By part (iii) of initial separability, conclude that

$$(z'_0, x_t, x_{t+1}, z'_{t+1}, \ldots) \succ (z'_0, y_t, y_{t+1}, z'_{t+1}, \ldots).$$

Apply quasi-stationarity $t - 1$ times to obtain

$$(z'_0, z'_1, \ldots, z'_{t-1}, x_t, x_{t+1}, z'_{t+1}, \ldots) \succ (z'_0, z'_1, \ldots, z'_{t-1}, y_t, y_{t+1}, z'_{t+1}, \ldots).$$

The proof of the independence of $\{t, t + 1, \ldots\}$ for $t = 2, \ldots$ is analogous.

Step 2. Show that any period t is sensitive. To see that, observe that by sensitivity of the period $t = 1$ there exists $x \in \mathcal{F}$ and $c, c' \in C$ such that

$$(x_0, c, x_{t+1}, x_{t+2}, \ldots) \succ (x_0, c', x_{t+1}, x_{t+2}, \ldots).$$

By quasi-stationarity, applied $t - 1$ times conclude that

$$(x_0, x_1, \ldots, x_{t-1}, c, x_{t+1}, x_{t+2}, \ldots) \succ (x_0, x_1, \ldots, x_{t-1}, c', x_{t+1}, x_{t+2}, \ldots).$$

Step 3. Additive representation on X_T. Fix $T \geq 1$ and fix $e \in C$. Weak order, finite continuity, and steps 1 and 2 imply that (by theorem 1 of Gorman 1968, together with Vind 1971) the restriction of $\succ$ to X_T is represented by

$$(x_0, x_1, \ldots, x_T, c, c, \ldots) \mapsto \sum_{t=0}^{T} v_{t,T}(x_t) + R_T(c),$$

for some nonconstant and continuous maps $v_{t,T}$ and R_T from C to $\mathbb{R}$. By the uniqueness of additive representations, the above functions can be chosen to satisfy

$$(5) \qquad\qquad v_{t,T}(e) = R_T(e) = 0.$$

Step 4. Since any $X_T \subseteq X_{T+1}$, there are two additive representations of $\succ$ on X_T:

$$(x_0, x_1, \ldots, x_T, c, c, \ldots) \mapsto \sum_{t=0}^{T} v_{t,T}(x_t) + R_T(c)$$

and

$$(x_0, x_1, \ldots, x_T, c, c, \ldots) \mapsto \sum_{t=0}^{T} v_{t,T+1}(x_t) + v_{T+1,T+1}(x_t) + R_{T+1}(c).$$

By the uniqueness of additive representations and the normalization (5), the above functions must satisfy $v_{t,T+1}(c) = \gamma_{T+1} v_{t,T}(c)$ for $t = 0, 1, \ldots, T - 1$ and $v_{T+1,T+1}(c) + R_{T+1}(c) = \gamma_{T+1} R_T(c)$ for some $\gamma_{T+1} > 0$. By the uniqueness of additive representations the representations can be normalized so that $\gamma_{T+1} = 1$. Let v_t denote the common function $v_{t,T}$. With this notation, we obtain

$$(6) \qquad\qquad v_{T+1}(c) + R_{T+1}(c) = R_T(c).$$

Step 5. By quasi-stationarity, for any $T \geq 1$ the two additive representations of $\succ$ on X_T:

$$(e, x_0, x_1, \ldots, x_{T-1}, c, c, \ldots) \mapsto v_0(e) + \sum_{t=1}^{T} v_t(x_{t-1}) + R_T(c)$$

and

$$(e, x_0, x_1, \ldots, x_{T-1}, c, c, \ldots) \mapsto v_0(e) + \sum_{t=1}^{T} v_{t+1}(x_{t-1}) + R_{T+1}(c),$$

represent the same preference. By the uniqueness of additive representations, and the normalization (5), there exists $\delta_T > 0$ such that for all $t = 1, 2, \ldots$, $v_{t+1}(c) = \delta_T v_t(c)$ for all $c \in C$ and $R_{T+1}(c) = \delta_T R_T(c)$. Note, that δ_T is independent of T, since the functions v and R are independent of T; let δ denote this common value.

Step 6. Define $u := v_0, v := \delta^{-1} v_1$ and $R := \delta^{-2} R_1$. With this notation, equation (6) is $\delta^{T+1} v(c) + \delta^{T+2} R(c) = \delta^{T+1} R(c)$ for all $c \in C$. Observe that $\delta = 1$ implies that v is a constant function, which is a contradiction; hence, $\delta \neq 1$. Thus, $R(c) = \frac{1}{1-\delta} v(c)$ for all $c \in C$. Thus, the preference on X_T is represented by

$$(x_0, x_1, \ldots, x_T, c, c, \ldots) \mapsto u(x_0) + \sum_{t=1}^{T} \delta^t v(x_t) + \frac{\delta^{T+1}}{1-\delta} v(c).$$

To rule out $\delta > 1$ note that because v is nonconstant, there exist $a, b \in C$ such that $v(a) > v(b)$. Then, since $\delta + \frac{\delta^2}{1-\delta} < 0$ it follows that $u(a) + \delta v(b) + \frac{\delta^2}{1-\delta} v(b) > u(a) + \delta v(a) + \frac{\delta^2}{1-\delta} v(a)$, so $eb \succ a$. However, by tail continuity there exists T such that $(eb)_T a \succ a$, which implies that

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

Thus, $(\delta + \cdots + \delta^T)(v(b) - v(a)) > 0$, which contradicts $v(a) > v(b)$ and $\delta > 0$. Thus, $\delta < 1$ and $U(x)$ represents $\succ$ on X_T for any T.

Step 7. Fix $x \in \mathcal{F}$. By constant equivalence, there exists $c \in C$ with $x \sim c$. Suppose there exists $a \in C$ such that $c \succ a$. Then by tail continuity there exists τ such that for all

$T \geq \tau$, $x_T a \succ a$, which by step 6 implies that $U(x_T a) > U(a)$. This implies that

$$\exists_\tau \forall_{T \geq \tau} u(x_0) + \sum_{t=1}^{T} \delta^t v(x_t) + \frac{\delta^{T+1}}{1-\delta} v(a) > u(a) + \sum_{t=1}^{T} \delta^t v(a) + \frac{\delta^{T+1}}{1-\delta} v(a)$$

$$\exists_\tau \forall_{T \geq \tau} \sum_{t=1}^{T} [\delta^t v(x_t) - \delta^t v(a)] > [u(a) - u(x_0)]$$

$$\exists_\tau \inf_{T \geq \tau} \sum_{t=1}^{T} [\delta^t v(x_t) - \delta^t v(a)] \geq [u(a) - u(x_0)]$$

$$\sup_\tau \inf_{T \geq \tau} \sum_{t=1}^{T} [\delta^t v(x_t) - \delta^t v(a)] \geq [u(a) - u(x_0)],$$

which means that $\liminf_T \sum_{t=1}^{T} +[\delta^t v(x_t) - \delta^t v(a)] \geq [u(a) - u(x_0)]$. Because the sequence $\sum_{t=1}^{T} \delta^t v(a)$ converges, it follows that

$$u(x_0) + \liminf_T \sum_{t=1}^{T} \delta^t v(x_t) \geq u(a) + \lim_T \sum_{t=1}^{T} \delta^t v(a) = U(a).$$

Since this is true for all $a \prec c$, by connectedness of C and continuity of u and v it follows that

$$(7) \qquad u(x_0) + \liminf_T \sum_{t=1}^{T} \delta^t v(x_t) \geq U(c).$$

On the other hand, suppose that $a \succsim c$ for all $a \in C$. Then, by constant equivalence for all T there exists $b \in C$ such that $x_T c \sim b$. This implies that $x_T c \succsim c$. Thus,

$$\forall_T u(x_0) + \sum_{t=1}^{T} \delta^t v(x_t) + \frac{\delta^{T+1}}{1-\delta} v(c) \geq u(c) + \sum_{t=1}^{T} \delta^t v(c) + \frac{\delta^{T+1}}{1-\delta} v(c)$$

$$\forall_T \sum_{t=1}^{T} \delta^t v(x_t) - \sum_{t=1}^{T} \delta^t v(c) \geq u(c) - u(x_0)$$

$$\liminf_T \sum_{t=1}^{T} \delta^t v(x_t) - \sum_{t=1}^{T} \delta^t v(c) \geq u(c) - u(x_0).$$

Since the sequence $\sum_{t=1}^{T} \delta^t v(c)$ converges, equation (7) follows.

An analogous argument implies that $\limsup_T \sum_{t=0}^{T} \delta^t v(x_t) \leq U(c)$, which establishes the existence of the limit of the partial sums and the representation. □

Proof of Theorem 2

We have

$$(e, b, a, \ldots) \succsim (e, a, b, \ldots)$$

iff

$$u(e) + \delta v(b) + \frac{\delta^2}{1-\delta} v(a) \geq u(e) + \delta v(a) + \frac{\delta^2}{1-\delta} v(b)$$

iff

$$v(b) + \frac{\delta}{1-\delta} v(a) \geq v(a) + \frac{\delta}{1-\delta} v(b)$$

iff

$$[v(b) - v(a)] \frac{1 - 2\delta}{1 - \delta} \geq 0$$

iff

$$1 - 2\delta \leq 0 \qquad\qquad □$$

Proof of Theorem 3

The following lemma is key in the proof of Theorem 3.

LEMMA 1. For any $\delta \in [0.5, 1]$ and any $\beta \in (0, 1]$ there exists a sequence $\{\alpha_t\}_t$ of elements in $\{0, 1\}$ such that $\beta = \sum_{t=0}^{\infty} \alpha_t \delta^t$.

Proof. Let $d_0 := 0$ and $\alpha_0 := 0$ and define the sequences $\{d_t\}$ and $\{\alpha_t\}$ by

$$d_{t+1} := \begin{cases} d_t + \delta^{t+1} & \text{if } d_t + \delta^{t+1} \leq \beta \\ d_t & \text{otherwise.} \end{cases}$$

and

$$\alpha_{t+1} := \begin{cases} 1 & \text{if } d_t + \delta^{t+1} \leq \beta \\ 0 & \text{otherwise.} \end{cases}$$

Since the sequence $\{d_n\}$ is increasing and bounded from above by β, it must converge; let $d := \lim d_t$. It follows that $d = \sum_{t=0}^{\infty} \alpha_t \delta^t$. Suppose that $d < \beta$. It follows that $\alpha_t = 1$ for almost all t; otherwise there would exist arbitrarily large t with $\alpha_t = 0$, and since $\delta^t < \beta - d$ for some such t that would contradict the construction of the sequence $\{d_t\}$. Let $T := \max\{t : \alpha_t = 0\}$. We have $d = d_{T-1} + \frac{\delta^{T+1}}{1-\delta} \leq \beta$. Since $\delta \geq 0.5$, it follows that $\delta^T \leq \frac{\delta^{T+1}}{1-\delta}$, so $d_{T-1} + \delta^T \leq \beta$, which contradicts the construction of the sequence $\{d_t\}$. □

The necessity of Axioms 1–9 follows from Theorems 1 and 2 and Lemma 1. Suppose that Axioms 1–9 hold. By Theorems 1 and 2 the preference is represented by equation (4) with $\delta \geq 0.5$. Normalize u and v so that there exists $\hat{e} \in C$ with $u(\hat{e}) = v(\hat{e}) = 0$. Let M be as in Axiom 9. Define $\gamma := \sum_{t \in M} \delta^{t-1}$. Axiom 9 implies that for all $a, b, c, d \in C$

$$u(a) + \delta v(b) > u(c) + \delta v(d)$$

if and only if

$$v(a) + \gamma v(b) > v(c) + \gamma v(d).$$

By the uniqueness of the additive representations, there exists $\beta > 0$ and $\lambda_1, \lambda_2 \in \mathbb{R}$ such that $v(e) = \beta u(e) + \lambda_1$ and $\gamma v(e) = \beta \delta v(e) + \lambda_2$ for all $e \in C$. By the above normalization, $\lambda_1 = \lambda_2 = 0$. Hence, $v(e) = \beta u(e)$ for all $e \in C$ and $\beta = \sum_{t \in M} \delta^{t-2}$. □

Proof of Theorem 4

The necessity of Axioms 1–7 and 10 is straightforward. For Axiom 11, if $(b, e_2) \succsim (a, e_1)$, $(c, e_1) \succsim (d, e_2)$ and $(e_3, a) \sim (e_4, b)$, it follows that:

$$(8) \qquad u(b) + \frac{\delta}{1-\delta} \beta u(e_2) \geq u(a) + \frac{\delta}{1-\delta} \beta u(e_1)$$

$$(9) \qquad u(c) + \frac{\delta}{1-\delta} \beta u(e_1) \geq u(d) + \frac{\delta}{1-\delta} \beta u(e_2)$$

$$(10) \qquad u(e_3) + \frac{\delta}{1-\delta} \beta u(a) = u(e_4) + \frac{\delta}{1-\delta} \beta u(b).$$

Equations (8)–(9) imply $u(b) - u(a) \geq u(d) - u(c)$. Suppose that the implication of Axiom 11 does not hold, so that $(e_4, d) \succ (e_3, c)$. Then

$$(11) \qquad u(e_4) + \frac{\delta}{1-\delta}\beta u(d) > u(e_3) + \frac{\delta}{1-\delta}\beta u(c).$$

Since $0 < \beta, 0 < \delta < 1$, equations (10)–(11) imply $u(d) - u(c) > u(b) - u(a)$. A contradiction. By analogy, the second condition of Axiom 11 is also necessary. Therefore, Axiom 11 is satisfied by the representation in Theorem 4.

Now we prove sufficiency. From Theorem 1 it follows that $\succsim$ admits the representation in equation (4). Define the binary relation $\succsim^*$ over the elements of C^2 as follows:

$$(b, c) \succsim^* (a, d)$$

$\Longleftrightarrow$ there exists $e_1, e_2, e_3, e_4 \in C$ such that

$$(12) \qquad (b, e_2) \succsim (a, e_1) \text{ and } (c, e_1) \succsim (d, e_2) \text{ and } (e_3, a) \sim (e_4, b).$$

We break the proof of sufficiency into four steps.

Step 1. First, we argue that $\succsim^*$ admits the following additive representation:

$$(b, c) \succsim^* (a, d) \Longleftrightarrow u(b) + u(c) \geq u(a) + u(d)$$

Using the definition of $\succsim^*$ and the representation (4) of $\succsim$, it follows that $(b, c) \succsim^* (a, d)$ implies the existence of elements $e_1, e_2 \in C$ such that:

$$u(a) + \frac{\delta}{1-\delta}v(e_1) \leq u(b) + \frac{\delta}{1-\delta}v(e_2)$$

and

$$u(d) + \frac{\delta}{1-\delta}v(e_2) \leq u(c) + \frac{\delta}{1-\delta}v(e_1).$$

Therefore $u(b) + u(c) \geq u(a) + u(d)$.

Now, suppose $u(b) + u(c) \geq u(a) + u(d)$. We consider the following six cases and we show that Condition 12 is satisfied.

   (i) $u(b) \geq u(a), u(c) \geq u(d), v(a) \geq v(b)$: Set $e = e_1 = e_2$ for any $e \in C$, and choose e_3, e_4 to satisfy $u(e_3) + \frac{\delta}{1-\delta}v(a) = u(e_4) + \frac{\delta}{1-\delta}v(b)$. Then, Condition (12) is satisfied.

(ii) $u(b) \geq u(a), u(c) \geq u(d), v(a) < v(b)$:   Set $e = e_1 = e_2$ for any $e \in C$ and choose e_3, e_4 to have $u(e_3) + \frac{\delta}{1-\delta} v(a) = u(e_4) + \frac{\delta}{1-\delta} v(b)$. Again, condition (12) is satisfied and $(b, c) \succsim^* (a, d)$.

(iii) $u(b) \geq u(a), u(c) < u(d), v(a) \geq v(b)$: Note that $u(b) - u(a) \geq u(d) - u(c) > 0$. Find e_1, e_2 to satisfy: $\frac{\delta}{1-\delta}[v(e_1) - v(e_2)] = u(d) - u(c) > 0$. Set $e = e_3, e_4$ to get indifference.

(iv) $u(b) \geq u(a), u(c) < u(d), v(a) < v(b)$: Do the same as above.

(v) $u(b) < u(a), u(c) \geq u(d), v(a) \geq v(b)$: Find e_1, e_2 to satisfy: $\frac{\delta}{1-\delta}[v(e_1) - v(e_2)] = u(b) - u(a) < 0$. Note that

$$
\begin{aligned}
0 = u(b) - u(a) &- \frac{\delta}{1-\delta}[v(e_1) - v(e_2)] \\
&\geq u(d) - u(c) - \frac{\delta}{1-\delta}[v(e_1) - v(e_2)].
\end{aligned}
$$

(vi) $u(b) < u(a), u(c) \geq u(d), v(a) < v(b)$: Do the same as above.

In any event $u(b) + u(c) \geq u(a) + u(d)$ implies $(b, c) \succsim^* (a, d)$. Therefore, the preference relation $\succsim^*$ admits an additive representation in terms of u.

Step 2. The preference relation $\succsim^*$ also admits a representation in terms of the index v:

$$
(b, c) \succsim^* (a, d) \Longleftrightarrow v(b) + v(c) \geq v(a) + v(d)
$$

Using the definition of $\succsim^*$ and Axiom 11 it follows that

$$
u(e_3) + \frac{\delta}{1-\delta} v(a) = u(e_4) + \frac{\delta}{1-\delta} v(b)
$$

and

$$
u(e_3) + \frac{\delta}{1-\delta} v(c) \geq u(e_4) + \frac{\delta}{1-\delta} v(d),
$$

which implies $v(b) + v(c) \geq v(a) + v(d)$. Now, for the other direction, we proceed as in step 1. Suppose $v(b) + v(c) \geq v(a) + v(d)$. Proceeding exactly as before, there are elements e_1, e_2, e_3, e_4 such that $(e_2, b) \succsim (e_1, a)$, $(e_1, c) \succsim (e_2, d)$ and $(a, e_3) \sim (b, e_4)$. By Axiom 11, it follows that $(c, e_3) \succsim (b, e_4)$. And therefore, $u(b) + u(c) \geq u(a) + u(d)$. Therefore, $(b, c) \succsim^* (a, d) \Longleftrightarrow v(b) + v(c) \geq v(a) + v(d)$.

Step 3. Since the preference relation $\succsim^*$ admits two different additive representations, it follows that the two utility indexes are related through a monotone affine transformation. This is, there exists $\beta > 0$ and γ such that for all $a \in C$:

$$v(a) = \beta u(a) + \gamma$$

We conclude that $\succsim$ is represented by the mapping

$$(13) \qquad x \mapsto u(x_0) + \beta \sum_{t=1}^{\infty} \delta^t u(x_t)$$

with $\beta > 0$.

Step 4. Take $a, c \in C$ such that $u(a) > u(c)$. The existence of such an element follows from the sensitivity axiom. Choose b, d to satisfy:

$$u(a) + \delta u(b) = u(c) + \delta u(d).$$

Axiom 10 implies that

$$u(a) + \beta \delta u(b) \geq u(c) + \beta \delta u(d).$$

The two inequalities imply $\beta \leq 1$.

Proof of Theorem 5

REMARK 1. Both Ghirardato and Marinacci (2001) and Nakamura (1990) study Choquet preferences, so their axioms have co-monotonicity requirements. To have simpler statements and to avoid introducing the concept of co-monotonicity in the main text, we use stronger axioms that hold for all, not necessarily co-monotone acts, but the co-monotone versions of those axioms could be used (are equivalent in the presence of other axioms).

The necessity of the axioms is straightforward. For sufficiency, we rely on the work of Ghirardato and Marinacci (2001). Note that their axiom B1 follows from our Axioms 1 and 2. Their axioms B2 and B3 follow from our Axiom 13. Their axiom S1 follows from the fact that by Theorem 1 the functions u and v are continuous. Finally their axiom S2 follows from our Axiom 12. Thus, by their lemma 31 there exists $\alpha \in (0, 1)$ and $w : C \to \mathbb{R}$

such that $(a, b) \mapsto \alpha w(a) + (1 - \alpha)w(b)$ represents $\succsim$. By uniqueness of additive representations, w is a positive affine transformation of u. Step iv in the proof of Theorem 5 concludes the proof. □

Nakamura's Axiom. An alternative to Theorem 5 is the following:

AXIOM 15. (Nakamura's A6) For $a, b, c, d \in C$ such that $b \succsim a$, $d \succsim c$, $d \succsim b$ and $c \succsim a$:

$$(c(a, b), c(c, d)) \sim (c(a, c), c(b, d))$$

and

$$(c(c, d), c(a, b)) \sim (c(c, a), c(d, b)).$$

THEOREM 7. The preference $\succsim$ satisfies Axioms 1–7 and 13–15 if and only if there exists a nonconstant and continuous function $u : C \to \mathbb{R}$ and parameters $\beta > 0$ and $\delta \in (0, 1)$ such that $\succsim$ is represented by the mapping

$$x \mapsto u(x_0) + \beta \sum_{t=1}^{\infty} \delta^t u(x_t).$$

Moreover, it satisfies Axiom 10 if and only if $\beta \leq 1$, that is, $\succsim$ has the quasi-hyperbolic discounting representation.

Proof. The necessity of Axioms 1–7, 10, and 13 is straightforward. For Axiom 15, take $a, b, c, d \in C$ as in the statement of the axiom and note that:

$$(14a) \qquad c(a, b) \equiv c_1, \quad u(c_1) + \frac{\delta}{1 - \delta}\beta u(c_1) = u(a) + \frac{\delta}{1 - \delta}\beta u(b)$$

$$(14b) \qquad c(c, d) \equiv c_2, \quad u(c_2) + \frac{\delta}{1 - \delta}\beta u(c_2) = u(c) + \frac{\delta}{1 - \delta}\beta u(d)$$

and also,

$$(15a) \qquad c(a, c) \equiv c_3, \quad u(c_3) + \frac{\delta}{1 - \delta}\beta u(c_3) = u(a) + \frac{\delta}{1 - \delta}\beta u(c)$$

$$(15b) \qquad c(b, d) \equiv c_4, \quad u(c_4) + \frac{\delta}{1 - \delta}\beta u(c_4) = u(b) + \frac{\delta}{1 - \delta}\beta u(d).$$

Therefore, using equations (14a)–(14b),

$$\left[1 + \beta\frac{\delta}{1-\delta}\right]\left[u(c_1) + \frac{\delta}{1-\delta}u(c_2)\right] = u(a) + \frac{\delta}{1-\delta}\beta u(b)$$
$$+ \frac{\delta}{1-\delta}\beta u(c) + \left(\frac{\delta}{1-\delta}\beta\right)^2 u(d)$$

(16)

and using equations (15a)–(15b),

$$\left[1 + \beta\frac{\delta}{1-\delta}\right]\left[u(c_3) + \frac{\delta}{1-\delta}u(c_4)\right] = u(a) + \frac{\delta}{1-\delta}\beta u(c)$$
$$+ \frac{\delta}{1-\delta}\beta u(b) + \left(\frac{\delta}{1-\delta}\beta\right)^2 u(d).$$

(17)

So, $(c_1, c_2) \sim (c_3, c_4)$. The second implication of Axiom 15 follows by analogy.

   For sufficiency of the axioms we rely on the proof of Lemma 3 (Proposition 1) in Nakamura (1990).[16] The argument goes as follows. Consider the restriction of $\succsim$ to elements of the form (a, b), with $a, b \in C$ and $b \succsim a$. Denote it by $\succsim_R$. The proof of Theorem 1 implies lemma 2 (parts 1 and 2) of Nakamura (1990), with $S = (s_1, s_2)$, $A = s_1$, $\phi \equiv u$ and $\psi \equiv \frac{\delta}{1-\delta}v$. Our Axioms 13 and 15 coincide exactly with A3 and A6 in Nakamura (1990) when $S = (s_1, s_2)$. Therefore, Lemma 3 implies there is a real valued function $r(x)$ such that:

$$(a, b) \succsim_R (c, d) \iff \alpha r(a) + (1-\alpha)r(b) \geq \alpha r(c) + (1-\alpha)r(d),$$

where r is defined (Nakamura 1990, p. 356) as $\frac{\phi(c)}{\alpha}$ for all $c \in C$ and $\alpha = \frac{1}{(1+\beta^*)}$, with β^* such that $\psi(c) = \beta^*\phi(c) + \gamma^*$, $\beta^* > 0$. Hence, it follows that for every $c \in C$, $\frac{\delta}{1-\delta}v(c) = \beta^*u(c) + \gamma^*$. If we set $\beta = \frac{1}{\beta^*}$, then we get $u(c) = \frac{\delta}{1-\delta}\beta u(c) + \gamma$. The representation (4) becomes:

$$x \mapsto u(x_0) + \beta\sum_{t=1}^{\infty}\delta^t u(x_t), \quad \beta > 0.$$

Step 4 in the proof of Theorem 5 concludes the proof.   □

   16. Nakamura's results are used explicitly by Chew and Karni (1994) and implicitly by Ghirardato and Marinacci (2001).

Proof of Theorem 6

Suppose that V is defined by the utility function $u : C \to \mathbb{R}$ and the sequence $1 = D_0 > D_1 > \cdots$ such that $\sum_{t=0}^{\infty} D_t < \infty$. Let V^n be a semi-hyperbolic preference defined by the same utility function and $D_t^n = D_t$ for $t = 0, 1, \ldots, n+1$ and $D_t^n = D_{n+1}\delta^{t-n}$ for $t > n+1$, where $\delta = D_1$.

For each n define the functions $W^n(x) = \sum_{t=0}^{n} D_t u(x_t)$, $R^n(x) = \sum_{t=n+1}^{\infty} D_t u(x_t)$, and $E^n(x) = D_{n+1} \sum_{t=n+1}^{\infty} \delta^{t-n-1} u(x_t)$. Notice that $V(x) = W^n(x) + R^n(x)$ for any n since the value of the sum is independent of n. Also, $V^n(x) = W^n(x) + E^n(x)$ for all n. Since the stream x is bounded, all these terms are well defined, and moreover the terms $E^n(x)$ and $R^n(x)$ converge to 0. Notice that this also implies that $D_{n+1} \to 0$. Suppose that there exist $\underline{u} < \overline{u}$ such that $\underline{u} \le u(x_t) \le \overline{u}$ for all t and define $M := \max\{|\overline{u}|, |\underline{u}|\}$. We have:

$$|V(x) - V^n(x)| = |W^n(x) + R^n(x) - W^n(x) - E^n(x)| = |R^n(x) - E^n(x)|$$

$$\le |R^n(x)| + |E^n(x)| \le M\left(\sum_{t=n+1}^{\infty} D_t + D_{n+1} \sum_{t=n+1}^{\infty} \delta^{t-n} \right) \to 0.$$

This also proves uniform convergence over all x within $\underline{u}, \overline{u}$.

Finally, notice that if $x \succsim^n y$ for n large enough, then $V^n(x) \ge V^n(y)$ for large n, so by the foregoing result $V(x) \ge V(y)$. Moreover, if for some $\epsilon > 0$ we have $V(x) - V(y) > \epsilon$ then since $V^n(x) \to V(x)$ and $V^n(y) \to V(y)$, we have $\lim_n [V^n(x) - V^n(y)] \ge \epsilon$, so $x \succ^n y$ for n sufficiently large. □

APPENDIX II: EMPIRICAL RESULTS

Econometric Analysis

Each agent i answers seven questions in each of the two price lists. We summarize each agent's set of answers by the "switch point" in each list; that is, we report the number of the first question (1 to 7) in which the agent chooses the impatient prospect A. If agent i always chooses the patient prospect B, we say that the switch point has a numerical value of 8. As noted before, under the assumption of quasi-hyperbolic discounting the agent has at most one switch point in each list, that is, she answers B for questions $1, \ldots, k$ and A for questions $k+1, \ldots, 7$.

Let $(s_{i,1}, s_{i,2})$ denote the switch points of agent i in price list 1 and 2, respectively. The objective of the econometric analysis in this article is to estimate the marginal distributions of (δ_i, β_i) in the population based on a sample of switch points for agents $i = 1, \ldots I$. In the following subsections we argue that our experimental design allows us to partially identify the marginal distributions of δ_i and β_i.

Data and Distributions of Switch Points. Our initial sample consists of two groups of subjects. The Money Group (M) has 639 subjects that answered the "money" questionnaire. The Ice Cream Group (IC) has 640 subjects that answered the "ice cream" questionnaire. We associate subjects with an IP address and we verify that there is no IP repetition inside the group. Consequently, we do not allow for a single IP address to answer the same questionnaire more than once.

We select a subsample of 336 subjects from the M group and 444 subjects from the IC group. The selection is based on three criteria (monotonicity, understanding, and consistency) described in the Online Appendix. For the selected sample, we focus on the distributions of switch points. These distributions are described in Figure A.I.

Our objective is to map the joint empirical distribution of switch points in Figure A.I into estimated lower and upper bounds for the marginal distributions of β_i and δ_i.

Marginal Distribution of δ_i. Partial identification: For $\delta \in [0, 1)$, let $F(\delta)$ denote the measure of the set of quasi-hyperbolic agents in the MTurk population (denoted $\mathcal{P}$) with parameter $\delta_i \leq \delta$. That is:

$$F(\delta) = \mu\{i \in \mathcal{P} \mid \delta_i \leq \delta\}.$$

We argue now that $F(\delta)$ is partially identified by the switch points in the second price list. Let $\delta^*(j)$ be the value of the discount factor that makes any agent i indifferent between options A and B in question j of the second price list, $j = 1 \ldots 7$. Note that $\delta^*(j)$ is defined by the equation:

$$\beta_i u_i(x) + \beta_i \delta^*(j)^{t_j} u_i(y) + \beta_i \delta^*(j)^{t_j+1} u_i(y)$$
$$= \beta_i u_i(y) + \beta_i \delta^*(j)^{t_j} u_i(x) + \beta_i \delta^*(j)^{t_j+1} u_i(x),$$

(**a**)

Empirical Distribution of Switch Points
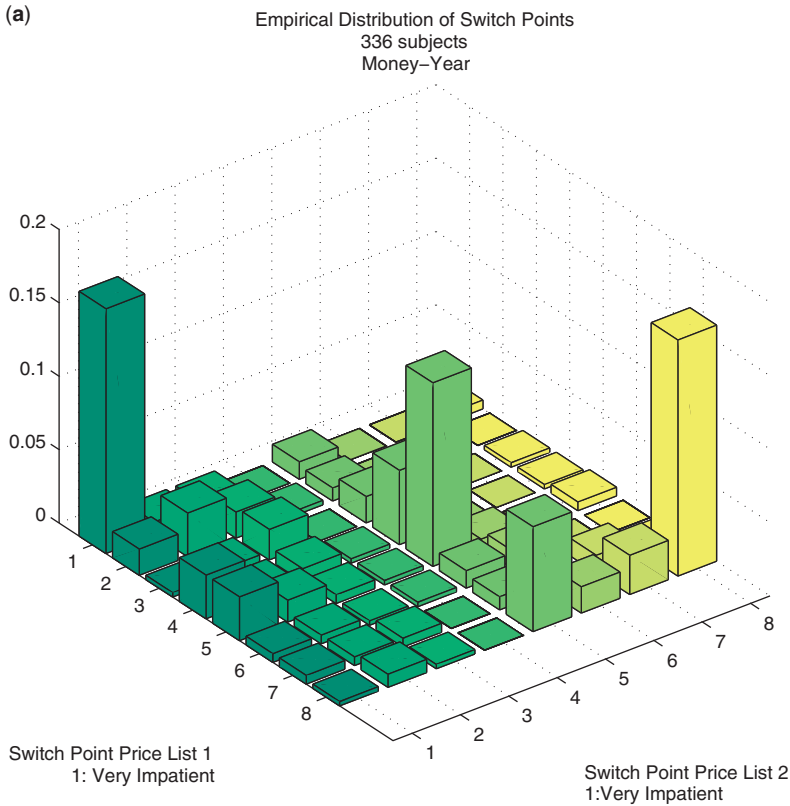336 subjects
Money–Year

FIGURE A.I

Distribution of Switch Points in the Sample

where $t_j = \{1, 3, 6, 12, 24, 36, 60\}$. If $u_i(x) > u_i(y)$ the latter holds if and only if:

$$(18) \qquad 1 = \delta^*(j)^{t_j} + \delta^*(j)^{t_j+1},$$

which has only one real solution in $[0, 1)$. The collection of intervals

$$[\delta^*(0), \delta^*(1)), [\delta^*(1), \delta^*(2)) \ldots [\delta^*(7), \delta^*(1))$$

is a partition of $[0, 1)$ (with $\delta^*(0) \equiv 0$ and $\delta^*(8) \equiv 1$).

(b)

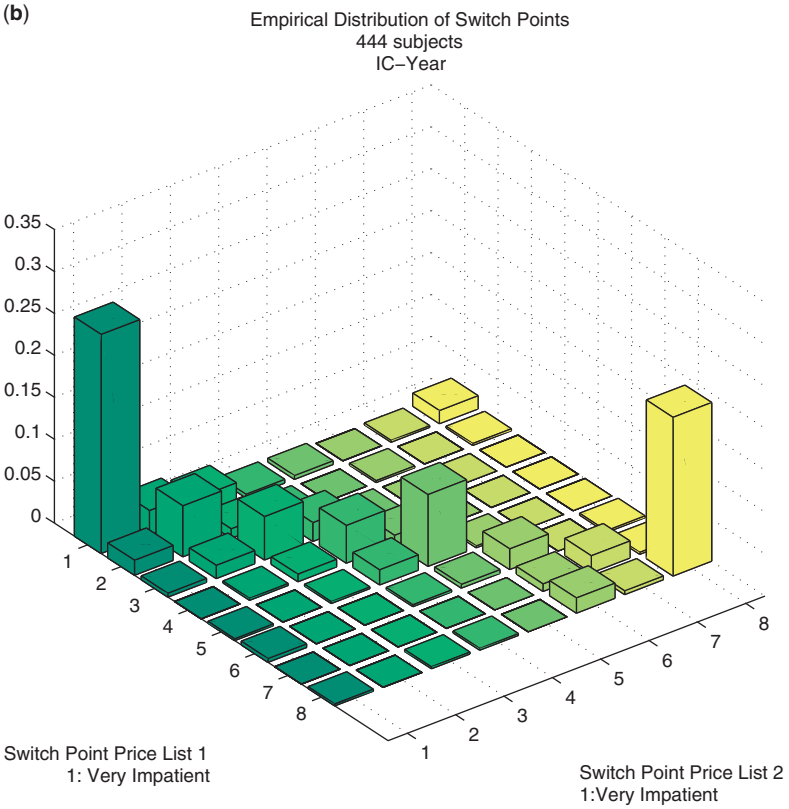Empirical Distribution of Switch Points
444 subjects
IC–Year

FIGURE A.I

Continued

PROPOSITION 1. For $j = 1 \ldots 7$

$$\mu\{i \in \mathcal{P} \mid s_{i,2} \leq j\} \leq F(\delta^*(j)) \leq \mu\{i \in \mathcal{P} \mid s_{i,2} \leq j + 1\}.$$

Proof. Note that

$$
\begin{aligned}
\mu\{i \in \mathcal{P} \mid s_{i,2} \leq j\} &= \mu\{i \in \mathcal{P} \mid i \text{ chooses A in question } j\} \\
&\leq \mu\{i \in \mathcal{P} \mid \delta_i^{t_j} + \delta_i^{t_j+1} \leq 1 = \delta^*(j)^{t_j} + \delta^*(j)^{t_j+1}\} \\
&= \mu\{i \in \mathcal{P} \mid \delta_i \leq \delta^*(j)\} \\
&= F(\delta^*(j)).
\end{aligned}
$$

Likewise,

$$F(\delta^*(j)) \le \mu\{i \in \mathcal{P} \,|\, \delta_i < \delta^*(j+1)\}$$
$$= \mu\{i \in \mathcal{P} \,|\, \delta_i^{t_{j+1}} + \delta_i^{t_{j+1}+1} < \delta^*(j+1)^{t_{j+1}} + \delta^*(j+1)^{t_{j+1}+1} = 1\}$$
$$\le \mu\{i \in \mathcal{P} \,|\, s_{i,2} \le j+1\} \qquad \square$$

COROLLARY. For any $\delta \in [\delta^*(j), \delta^*(j+1))$, $j = 1, \ldots 7$,

$$\underline{F}(\delta) \equiv \mu\{i \in \mathcal{P} \,|\, s_{i,2} \le j\} \le F(\delta) \le \mu\{i \in \mathcal{P} \,|\, s_{i,2} \le j+1\} \equiv \overline{F}(\delta).$$

Proof. For the lower bound, the weak monotonicity of the c.d.f. implies

$$F(\delta) \ge F(\delta^*(j))$$
$$\ge \mu\{i \in \mathcal{P} \,|\, s_{i,2} \le j\} \quad \text{(by Proposition 1)}.$$

For the upper bound:

$$F(\delta) \le \mu\{i \in \mathcal{P} \,|\, \delta_i < \delta^*(j+1)\}$$
$$\le \mu\{i \in \mathcal{P} \,|\, s_{i,2} \le j+1\} \qquad \square$$

Hence, the marginal distribution of δ_i is partially identified by the switch points $s_{i,2}$.

Estimation and inference: lower and upper bounds. Our inference problem falls in the set-up considered by Imbens and Manski (2004) and Stoye (2009): a real-valued parameter, $F(\delta)$, is partially identified by an interval whose upper and lower bounds may be estimated from sample data. Given the results in Proposition 1 and its corollary, we consider the following estimators for the lower and upper bounds of $F(\delta)$. For any $\delta \in [\delta^*(j), \delta^*(j+1)]$:

$$\hat{\underline{F}}(\delta) \equiv \frac{1}{I} \sum_{i=1}^{I} 1\{s_{i,2} \le j\}$$

and

$$\hat{\overline{F}}(\delta) \equiv \frac{1}{I} \sum_{i=1}^{I} 1\{s_{i,2} \le j+1\}$$
$$= \hat{\underline{F}}(\delta) + \frac{1}{I} \sum_{i=1}^{I} \{s_{i,2} = j+1\}.$$

If the preference parameters (δ_i, β_i) are independent draws from the distribution μ, then the weak law of large numbers implies that:

$$\hat{\underline{F}}(\delta) \xrightarrow{p} \underline{F}(\delta) \quad \text{and} \quad \hat{\overline{F}}(\delta) \xrightarrow{p} \overline{F}(\delta).$$

To construct confidence bands for the partially identified parameter we use Imbens and Manski (2004)'s approach as described in Stoye (2009, p. 1301). For each δ we consider a confidence set for the parameter $F(\delta^*(1)) \le F(\delta) \le F(\delta^*(7))$ of the form:

$$(19) \qquad CI_\alpha \equiv \left[\hat{\underline{F}}(\delta) - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{I}}, \hat{\overline{F}}(\delta) + \frac{c_\alpha \hat{\sigma}_u}{\sqrt{I}} \right]$$

where

$$\hat{\sigma}_l = \left(\hat{\underline{F}}(\delta)(1 - \hat{\underline{F}}(\delta)) \right)^{1/2} \quad \text{and} \quad \hat{\sigma}_u = \left(\hat{\overline{F}}(\delta)(1 - \hat{\overline{F}}(\delta)) \right)^{1/2}$$

and c_α satisfies

$$\Phi\left(c_\alpha + \frac{\sqrt{I}\hat{\Delta}}{\max\{\hat{\sigma}_l, \hat{\sigma}_u\}}\right) - \Phi(-c_\alpha) = 1 - \alpha,$$

$$\hat{\Delta} = \hat{\overline{F}}(\delta) - \hat{\underline{F}}(\delta) = \frac{1}{I}\sum_{i=1}^{I} \{s_{i,2} = j + 1\}.$$

Figure A.II shows the estimated upper and lower bounds and the (point-wise) confidence sets for $F(\delta)$. Each of the jumps of bounds for the c.d.f. occurs at the (real) roots of the equations

$$1 = \delta^*(j)^{t_j} + \delta^*(j)^{t_j+1},$$

where t_j corresponds to the delay of the rewards in the second price list. So, based on our experimental design, the seven jumps for the bounds of the c.d.f. occur at:

$$\delta^*(1) = 0.6180, \quad \delta^*(2) = 0.8192, \quad \delta^*(3) = 0.8987, \quad \delta^*(4) = 0.9460$$
$$\delta^*(5) = 0.9721, \quad \delta^*(6) = 0.9812, \quad \delta^*(7) = 0.9886.$$

Marginal Distribution of β_i. For $\beta \ge 0$, let $G(\beta)$ denote the measure of the set of quasi-hyperbolic agents in the MTurk
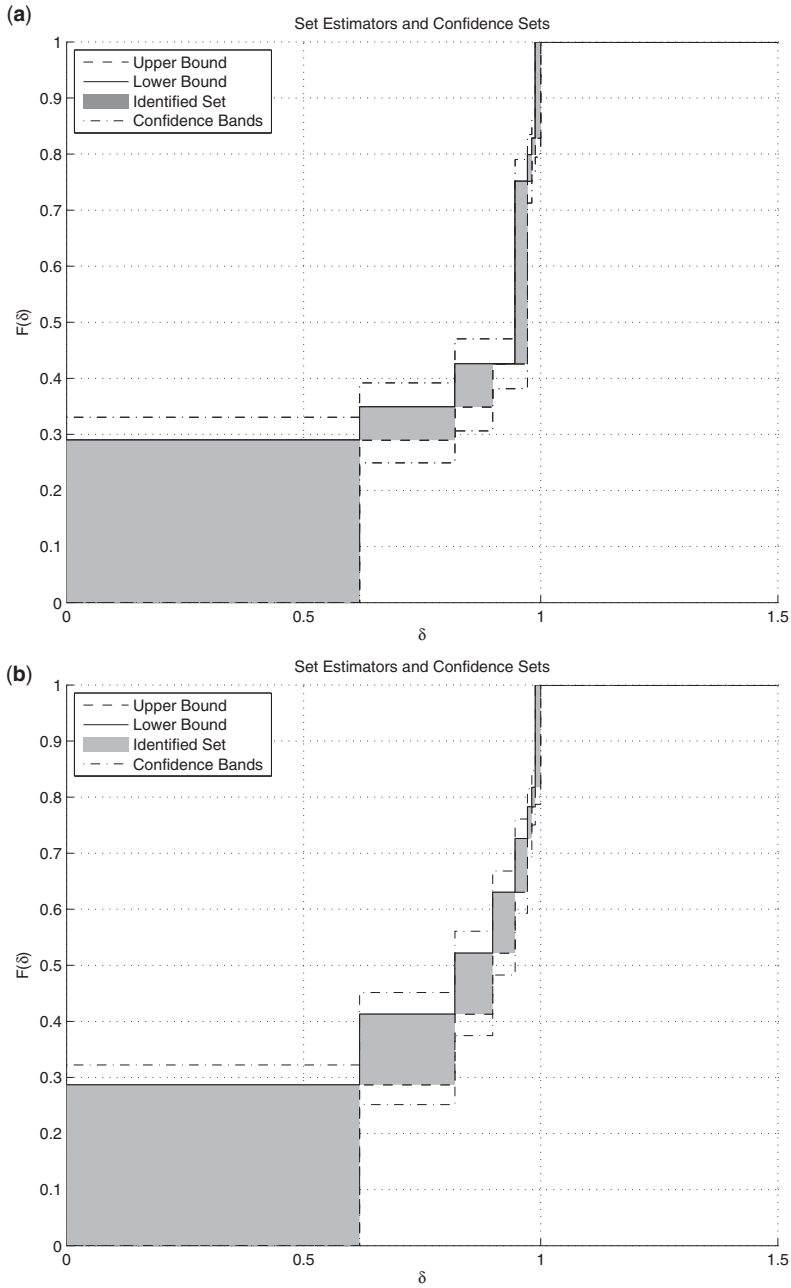
(a)

(b)

FIGURE A.II

Bounds for $F(\delta)$: a) Money b) Ice cream

population (denoted $\mathcal{P}$) with parameter $0 \leq \beta_i \leq \beta$. That is:

$$G(\beta) \equiv \mu\{i \in \mathcal{P} | \beta_i \leq \beta\}.$$

We now show that the switch points in the first and second price lists allows us to partially identify the marginal distribution $G(\beta)$. Let $\delta^*(j)$ be the solution to equation (18). For $j = 1, 2 \ldots 7$ and $k = 1, 2, \ldots 7$, define:

$$\beta^*(j, k) \equiv \frac{1}{\delta^*(j)^{t_k} + \delta^*(j)^{t_k+1}},$$

where $t_k = \{1, 3, 6, 12, 24, 36, 60\}$. Note that t_k represents the first future payment date in questions A and B of price list 1. Define:

$$(20) \qquad \underline{n}(j \,|\, \beta) \equiv \max\{n \,|\, \beta^*(j, n) \leq \beta\}$$

$$(21) \qquad \overline{n}(j \,|\, \beta) \equiv \min\{n \,|\, \beta < \beta^*(j+1, n)\}.$$

Let $\underline{n}(0 \,|\, \beta) \equiv 0$ for all β. We start by proving the following result:

LEMMA 2. For $j = 0, \ldots 7$, $\beta \geq 0$, let

$$B(j, \beta) = \{i \in \mathcal{P} \,|\, 0 \leq \beta_i \leq \beta, \quad s_{i,2} = j+1\}.$$

$$\{i \in \mathcal{P} \,|\, s_{i,1} \leq \underline{n}(j \,|\, \beta), \quad s_{i,2} = j+1\} \subseteq B(j, \beta)$$
$$(22) \qquad\qquad\qquad \subseteq \{i \in \mathcal{P} \,|\, s_{i,1} \leq \overline{n}(j \,|\, \beta), \quad s_{i,2} = j+1\}$$

Proof. We establish the lower bound first. The result holds for vacuously for $j = 0$. So, suppose $j > 0$. Note that $s_{i,2} = j+1$ implies two things. First, the switch point in the second price list did not occur at $j < j+1$. Therefore,

$$1 \leq \delta_i^{t_j} + \delta_i^{t_j+1},$$

where t_j corresponds to the first future payment date in question j of price list 2. By definition of $\delta^*(j)$, the latter implies

$$\delta^*(j)^{t_j} + \delta^*(j)^{t_j+1} \leq \delta_i^{t_j} + \delta_i^{t_j+1},$$

which implies $\delta_i \geq \delta^*(j)$. Second, at question $j+1$ the switch occurs. Hence:

$$\delta^*(j+1)^{t_{j+1}} + \delta^*(j+1)^{t_{j+1}+1} = 1 \geq \delta_i^{t_{j+1}} + \delta_i^{t_{j+1}+1}.$$

Consequently, $\delta^*(j+1) \geq \delta_i$. We conclude that for any i such that $s_{i,2} = j+1$:

$$(23) \qquad\qquad \delta_i \in [\delta^*(j), \delta^*(j+1)].$$

In addition, let $k' \leq \underline{n}(j \mid \beta)$. Note that for a quasi-hyperbolic agent $s_{i,1} = k$ implies

$$(24) \qquad \beta_i \leq \frac{1}{\delta_i^{t_k} + \delta_i^{t_k+1}} \leq \frac{1}{\delta_i^{t_{\underline{n}(j|\beta)}} + \delta_i^{t_{\underline{n}(j|\beta)}+1}} = \beta^*(j, \underline{n}(j \mid \beta)).$$

Hence $s_{i,1} \leq \underline{n}(j \mid \beta)$ and $s_{1,2} = j+1$ imply equations (23) and (24). Equation (20) implies

$$0 \leq \beta_i \leq \beta^*(j, \underline{n}(j \mid \beta)) \leq \beta,$$

and we conclude

$$\{i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), \quad s_{i,2} = j+1\} \subseteq B(j, \beta).$$

Now we establish the upper bound. Suppose $i \in B(j, \beta)$. Then i belongs to

$$B(j, \beta) \equiv \left\{i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta < \beta^*(j+1, \overline{n}(j+1 \mid \beta), \quad s_{1,2} = j+1\right\}.$$

Since

$$\begin{aligned} \beta_i < \beta^*(j+1, \overline{n}(j+1 \mid \beta) &= \frac{1}{\delta^*(j+1)^{t_{\overline{n}(j+1|\beta)}} + \delta^*(j+1)^{t_{\overline{n}(j+1|\beta)}+1}} \\ &\leq \frac{1}{\delta_i^{t_{\overline{n}(j+1|\beta)}} + \delta_i^{t_{\overline{n}(j+1|\beta)}+1}} \end{aligned}$$

the switch in price list 1 occurred at most at period $\overline{n}(j+1 \mid \beta)$. Therefore, $s_{i,1} \leq \overline{n}(j+1 \mid \beta)$. $\qquad \square$

We use the previous lemma to partially identify $G(\beta)$.

PROPOSITION 2 (BOUNDS FOR $G(\beta)$). For $j = 0, \ldots 7$:

(i) $\sum_{j=0}^{7} \mu\{i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), \quad s_{i,2} = j+1\} \leq G(\beta)$

(ii) $G(\beta) \leq \sum_{j=0}^{7} \mu\{i \in \mathcal{P} \mid s_{i,1} \leq \overline{n}(j \mid \beta), \ s_{i,2} = j+1\}.$

Proof. First we establish the lower bound. By Lemma 2, for each $j = 0, \ldots 7$:

$$\left\{i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), \ s_{i,2} = j+1\right\} \subseteq B(j, \beta)$$

Therefore,

$$\bigcup_{j=0}^{7} \left\{ i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), s_{i,2} = j+1 \right\} \subseteq \bigcup_{j=0}^{7} B(j, \beta)$$

$$= \bigcup_{j=0}^{7} \left\{ i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta^*(j, \underline{n}(j \mid \beta)), \ s_{i,2} = j+1 \right\}$$

$$\subseteq \bigcup_{j=0}^{7} \left\{ i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta, \ s_{i,2} = j+1 \right\}$$

$$= \{ i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta \}.$$

Hence,

$$\mu \left(\bigcup_{j=0}^{7} \{ i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), \ s_{i,2} = j+1 \} \right) \leq \bigcup_{j=0}^{7} \mu \left\{ i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta \right\}$$

$$= G(\beta).$$

Now we establish the upper bound. From Lemma 2:

$$\left\{ i \in \mathcal{P} \mid 0 \leq \beta_i \leq \beta, s_{i,2} = j+1 \right\}$$

is a subset of

$$\left\{ i \in \mathcal{P} \mid s_{i,1} \leq \overline{n}(j \mid \beta), s_{i,2} = j+1 \right\}.$$

The result then follows. □

Estimation and inference: lower and upper bounds. Based on Proposition 2, the estimators for the upper and lower bounds of the population are given by:

(i) $\sum_{j=0}^{7} \frac{1}{I} \sum_{i=1}^{I} 1 \left\{ i \in \mathcal{P} \mid s_{i,1} \leq \underline{n}(j \mid \beta), \ s_{i,2} = j+1 \right\}$

(ii) $\sum_{j=0}^{7} \frac{1}{I} \sum_{i=1}^{I} 1 \left\{ i \in \mathcal{P} \mid s_{i,1} \leq \overline{n}(j \mid \beta), \ s_{i,2} = j+1 \right\},$

which can be written as:

(i) $\underline{\hat{G}}(\beta) = \frac{1}{I} \sum_{i=1}^{I} 1 \left\{ i \in \mathcal{P} \mid \bigcup_{j=0}^{7} (s_{i,1} \leq \underline{n}(j \mid \beta), \ s_{i,2} = j+1) \right\}$

(ii) $\overline{\hat{G}}(\beta) = \frac{1}{I} \sum_{i=1}^{I} 1 \left\{ i \in \mathcal{P} \mid \bigcup_{j=0}^{7} (s_{i,1} \leq \overline{n}(j \mid \beta), \ s_{i,2} = j+1) \right\}.$
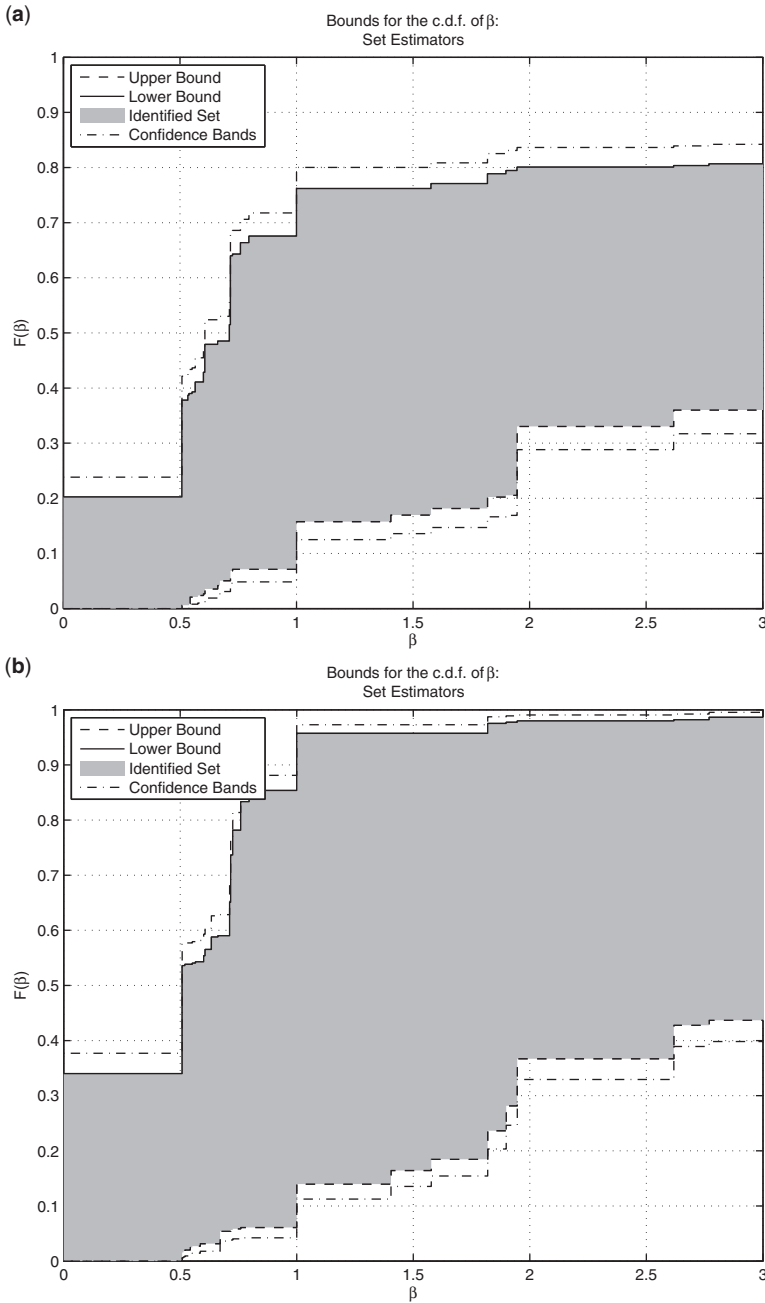
(a)

(b)

FIGURE A.III

Bounds for $G(\beta)$: a) Money b) Ice cream

Imbens and Manski (2004)'s approach is used to build a confidence set for the parameter $G(\beta)$:

$$(25) \qquad CI_\alpha \equiv \left[\underline{\hat{G}}(\beta) - \frac{c_\alpha \hat{\sigma}_l}{\sqrt{I}}, \ \overline{\hat{G}}(\beta) + \frac{c_\alpha \hat{\sigma}_u}{\sqrt{I}} \right],$$

where

$$\hat{\sigma}_l = \left(\underline{\hat{G}}(\beta)(1 - \underline{\hat{G}}(\delta)) \right)^{1/2} \quad \text{and} \quad \hat{\sigma}_u = \left(\overline{\hat{G}}(\beta) \, 1 - \overline{\hat{G}}(\beta)) \right)^{1/2},$$

and c_α satisfies

$$\Phi\left(c_\alpha + \frac{\sqrt{I}\hat{\Delta}}{\max\{\hat{\sigma}_l, \hat{\sigma}_u\}} \right) - \Phi(-c_\alpha) = 1 - \alpha,$$

$$\hat{\Delta} = \overline{\hat{G}}(\beta) - \underline{\hat{G}}(\beta).$$

Figure A.III reports the estimates for the lower and upper bounds along with a 95% confidence set for the partially identified parameter $G(\beta)$.

New York University
Harvard University

Supplementary Material

An Online Appendix for this article can be found at QJE online (qje.oxfordjournals.org).

References

Abdellaoui, Mohammed, Arthur E. Attema, and Han Bleichrodt, "Intertemporal Tradeoffs for Gains and Losses: An Experimental Measurement of Discounted Utility," *Economic Journal*, 120 (2010), 845–866.

Andersen, Steffen, Glenn W. Harrison, Morten I. Lau, and E. Elisabet Rutström, "Eliciting Risk and Time Preferences," *Econometrica*, 76 (2008), 583–618.

Andreoni, James, and Charles Sprenger, "Estimating Time Preferences from Convex Budgets," *American Economic Review*, 102 (2012), 3333–3356.

Andreoni, James, Michael A. Kuhn, and Charles Sprenger, "On Measuring Time Preferences," Working paper, UC San Diego, 2013.

Attema, A. E., H. Bleichrodt, K. I. M. Rohde, and P. P. Wakker, "Time-Tradeoff Sequences for Analyzing Discounting and Time Inconsistency," *Management Science*, 56 (2010), 2015–2030.

Augenblick, Ned, Muriel Niederle, and Charles Sprenger, "Working over Time: Dynamic Inconsistency in Real Effort Tasks," NBER Working Paper, 2013.

Bleichrodt, H., K. I. M. Rohde, and P. P. Wakker, "Koopmans' Constant Discounting for Intertemporal Choice: A Simplification and a Generalization," *Journal of Mathematical Psychology*, 52 (2008), 341–347.

Casari, M., and D. Dragone, "On Negative Time Preferences," *Economics Letters*, 111 (2010), 37–39.

Chabris, Christopher, David Laibson, and Jonathan Schuldt, "Intertemporal Choice," in *Palgrave Dictionary of Economics* (London: Palgrave, 2008).

Chew, H. C., and E. Karni, "Choquet Expected Utility with a Finite State Space: Commutativity and Act-Independence," *Journal of Economic Theory*, 62 (1994), 469–479.

Cohen, Michèle, Jean-Marc Tallon, and Jean-Christophe Vergnaud, "An Experimental Investigation of Imprecision Attitude and its Relation with Risk Attitude and Impatience," *Theory and Decision*, 71 (2011), 81–109.

Coller, Maribeth, and Melonie B. Williams, "Eliciting Individual Discount Rates," *Experimental Economics*, 2 (1999), 107–127.

Dohmen, Thomas J., Armin Falk, David Huffman, and Uwe Sunde, "Interpreting Time Horizon Effects in Inter-Temporal Choice," CESifo Working Paper Series No. 3750, 2012.

Epstein, Larry G., and S. E. Zin, "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework," *Econometrica*, 57 (1989), 937–969.

Fels, S., and R. Zeckhauser, "Perfect and Total Altruism across the Generations," *Journal of Risk and Uncertainty*, 37 (2008), 187–197.

Fishburn, P. C., and A. Rubinstein, "Time Preference," *International Economic Review*, 23 (1982), 677–694.

Fudenberg, Drew, and David K. Levine, "A Dual-Self Model of Impulse Control," *American Economic Review*, 96 (2006), 1449–1476.

Fudenberg, Drew, and Eric Maskin, "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, 53 (1991), 428–438.

Ghirardato, P., and M. Marinacci, "Risk, Ambiguity, and the Separation of Utility and Beliefs," *Mathematics of Operations Research*, 26 (2001), 864–890.

Gigliotti, Gary, and Barry Sopher, "Analysis of Intertemporal Choice: A New Framework and Experimental Results," *Theory and Decision*, 55 (2003), 209–233.

Gorman, W. M., "The Structure of Utility Functions," *Review of Economic Studies*, 35 (1968), 367–390.

Gul, Faruk, and Wolfgang Pesendorfer, "Temptation and Self-Control," *Econometrica*, 69 (2001), 1403–1435.

———, "Self-Control and the Theory of Consumption," *Econometrica*, 72 (2004), 119–158.

Halevy, Yoram, "Time Consistency: Stationarity and Time Invariance," Micro Theory Working Papers, 2012.

Harrison, Glenn W., Morten I. Lau, and Melonie B. Williams, "Estimating Individual Discount Rates in Denmark: A Field Experiment," *American Economic Review*, 92 (2002), 1606–1617.

Hayashi, T., "Quasi-Stationary Cardinal Utility and Present Bias," *Journal of Economic Theory*, 112 (2003), 343–352.

Horton, John J., David G. Rand, and Richard J. Zeckhauser, "The Online Laboratory: Conducting Experiments in a Real Labor Market," *Experimental Economics*, 14 (2011), 399–425.

Imbens, Guido W., and Charles F. Manski, "Confidence Intervals for Partially Identified Parameters," *Econometrica*, 72 (2004), 1845–1857.

Johnson, Matthew W., and Warren K. Bickel, "Within-Subject Comparison of Real and Hypothetical Money Rewards in Delay Discounting," *Journal of the Experimental Analysis of Behavior*, 77 (2002), 129–146.

Kochov, Asen, "Geometric Discounting in Discrete, Infinite-Horizon Choice Problems," Mimeo, 2013.

Koopmans, T. C., "Stationary Ordinal Utility and Impatience," *Econometrica*, 28 (1960), 287–309.

———, "Representation of Preference Orderings over Time, in *Decision and Organization*, C. B. McGuire, and Roy Radner eds. (Amsterdam: North-Holland, 1972).

Kreps, D. M., and E. L. Porteus, "Temporal Resolution of Uncertainty and Dynamic Choice Theory," *Econometrica*, 46 (1978), 185–200.

Laibson, D., "Golden Eggs and Hyperbolic Discounting," *Quarterly Journal of Economics*, 112 (1997), 443–477.

Loewenstein, G., and D. Prelec, "Anomalies in Intertemporal Choice: Evidence and an Interpretation," *Quarterly Journal of Economics*, 107 (1992), 573–597.

Luce, R. D., and J. W. Tukey, "Simultaneous Conjoint Measurement: A New Type of Fundamental Measurement," *Journal of Mathematical Psychology* (1964), 1–27.

Marge, Matthew, Satanjeev Banerjee, and Alexander I. Rudnicky, "Using the Amazon Mechanical Turk for Transcription of Spoken Language," Carnegie Mellon University Computer Science Department Paper 1350, 2010.

Mason, Winter, and Duncan J. Watts, "Financial Incentives and the Performance of Crowds," *ACM SigKDD Explorations Newsletter*, 11 (2010), 100–108.

Mason, Winter, and Siddharth Suri, "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavior Research Methods*, 44 (2012), 1–23.

McClure, S. M., K. M. Ericson, D. I. Laibson, G. Loewenstein, and J. D. Cohen, "Time Discounting for Primary Rewards," *Journal of Neuroscience*, 27 (2007), 5796–5804.

Nakamura, Y., "Subjective Expected Utility with Non-additive Probabilities on Finite State Spaces," *Journal of Economic Theory*, 51 (1990), 346–366.

Noor, J., "Hyperbolic Discounting and the Standard Model: Eliciting Discount Functions," *Journal of Economic Theory*, 144 (2009), 2077–2083.

———, "Temptation and Revealed Preference," *Econometrica*, 79 (2011), 601–644.

O'Donoghue, Ted, and Matthew Rabin, "Choice and Procrastination," *Quarterly Journal of Economics*, 116 (2001), 121–160.

Paolacci, Gabriele, Jesse Chandler, and Panagiotis Ipeirotis, "Running Experiments on Amazon Mechanical Turk," *Judgment and Decision Making*, 5 (2010), 411–419.

Phelps, E. S., and R. A. Pollak, "On Second-Best National Saving and Game-Equilibrium Growth," *Review of Economic Studies*, 35 (1968), 185–199.

Ramsey, F., "Truth and Probability, in Ramsey, *The Foundations of Mathematics and Other Logical Essays*, R. B. Braithwaite, ed. (London: Kegan Paul, Trench, Trubner, 1931), pp. 156–198.

Read, Daniel, "Is Time-Discounting Hyperbolic or Subadditive?," *Journal of Risk and Uncertainty*, 23 (2001), 5–32.

Samuelson, P. A., "A Note on Measurement of Utility," *Review of Economic Studies*, 4 (1937), 155–161.

Sayman, Serdar, and Ayse Öncüler, "An Investigation of Time Inconsistency," *Management Science*, 55 (2009), 470–482.

Scholten, Marc, and Daniel Read, "Discounting by Intervals: A Generalized Model of Intertemporal Choice," *Management Science*, 52 (2006), 1424–1436.

Sorin, S., "On Repeated Games with Complete Information," *Mathematics of Operations Research*, 11 (1986), 147–160.

Stoye, Jörg, "More on Confidence Intervals for Partially Identified Parameters," *Econometrica*, 77 (2009), 1299–1315.

Strotz, Robert Henry, "Myopia and Inconsistency in Dynamic Utility Maximization," *Review of Economic Studies*, 23 (1955), 165–180.

Takeuchi, Kan, "Non-parametric Test of Time Consistency: Present Bias and Future Bias," *Games and Economic Behavior*, 71 (2011), 456–478.

Thaler, R., "Some Empirical Evidence on Dynamic Inconsistency," *Economics Letters*, 8 (1981), 201–207.

Thaler, Richard H., and Hersh M. Shefrin, "An Economic Theory of Self-Control," *Journal of Political Economy*, 89 (1981), 392–406.

Van der Pol, Marjon, and John Cairns, "Descriptive Validity of Alternative Intertemporal Models for Health Outcomes: An Axiomatic Test," *Health Economics*, 20 (2011), 770–782.

Vind, K., "Note on 'The Structure of Utility Functions'," *Review of Economic Studies*, 38 (1971), 113.

Zeckhauser, R., and S. Fels, "Discounting for Proximity with Perfect and Total Altruism," Harvard Institute of Economic Research, Discussion Paper 50, 1968.