

Discounting and altruism to future decision-makers*

Maria Saez-Marti[†]

The Research Institute of Industrial Economics
Stockholm, Sweden
and

Jörgen W. Weibull[‡]

Department of Economics
Boston University

Abstract

Is discounting of future decision-makers' consumption utilities consistent with "pure" altruism toward those decision-makers, that is, a concern that they are better off according to their own, likewise forward-looking, preferences? It turns out that the answer is positive for many but not all discount functions used in the economics literature. In particular, "hyperbolic" discounting of the form used by Phelps and Pollak (1968) and Laibson (1997) is consistent with exponential altruism towards all future generations. More generally, we establish a one-to-one relationship between discount functions and altruism weight systems, and provide sufficient, as well as necessary, conditions for discount functions to be consistent with pure altruism.

JEL codes: D11, D64, D91, E21.

Keywords: Altruism, discounting, time preferences.

*This is a revised version of IUI Working Papers No. 572, 2002, "Welfare foundations of discounting" and No. 575, 2002, "Discounting and future selves." The authors thank Philippe Aghion, Cedric Argenton, Geir Asheim, John Leahy, Debraj Ray, Arthur Robson, Ariel Rubinstein and Fabrizio Zilibotti for helpful comments and suggestions. We are particularly grateful to Ulf Persson who proved and generalized one of our conjectures.

[†]Saez-Marti thanks Vetenskapsrådet for financial support of her research.

[‡]Weibull's research was partly carried out at the Stockholm School of Economics and at The Research Institute of Industrial Economics, Stockholm. Weibull thanks the Laboratoire d'Econométrie, Ecole Polytechnique, Paris, for its hospitality during part of his research.

1 Introduction

Many economics issues concern sequences of decisions. In models of such situations, the successive decisions are usually taken either by one and the same individual or by successive decision-makers, such as the generations in a dynasty. We here analyze a widely used class of time preferences in such analyses, namely those that permit utility representation as a sum of discounted instantaneous utilities. Hence, the decision-maker in each period has some concern for the future - be it his or her own, or that of future decision-makers. One may then ask the question why the current decision-maker is only concerned with the instantaneous utility of these future decision makers, when they, like the present decision-maker, also care about the future? We here ask whether discounting of future instantaneous utilities is consistent with “pure” altruism towards these future decision-makers, that is, a concern that future decision-makers are better off in terms of their own preferences (total utility).¹

Recognition of the economic importance of altruism goes back at least to Edgeworth (1881), who examined the effects of pure altruism on the contract curve in a two-person exchange economy. With X and Y denoting the two persons in question, Edgeworth wrote that

“we might suppose that the object which X (whose own utility is P), tends - in a calm, effective moment - to maximize, is not P , but $P + \lambda\Pi$; where λ is a *coefficient of effective sympathy*. And similarly Y - not of course while rushing to self-gratification, but in those regnant moments which characterize an ethical ‘method’ - may propose to himself as an end $\Pi + \mu P$.” (op. cit. p. 53).²

Many economists have analyzed, in a wide range of settings, the effect of altruism for economic agents’ decision-making. For instance, Barro’s (1974) famous analysis of Ricardian tax neutrality leans heavily on pure altruism of the type studied here: each generation cares about the next generation’s total utility, which in turn depends on the following generation’s total utility, in an infinite chain (see Bernheim (1987) for a survey of this literature, and see Andreoni (1989) for a model that allows for both pure and impure altruism). The present analysis identifies conditions under which a concern for future generations’ consumption utilities is, or is not, behaviorally equivalent to pure altruism towards these future generations. In his study of pure intergenerational altruism, Ray (1987) called for precisely such an investigation, and we hope to shed some light on this issue.³

¹By contrast, a decision-maker A is sometimes called “paternalistically” or “impurely” altruistic if A cares about others’ consumption, and/or directly about A ’s gifts or bequests to others (“warm glow effects”), without full regard to all factors of relevance for others’ well-being. Ray (1987) and Hori (2001) use the term “paternalistic altruism,” while Andreoni (1989) use the term “impure altruism.”

²See Collard (1975) for an analysis of Edgeworth’s treatment of altruism.

³“The representation of non-paternalistic functions in paternalistic form has ... been the subject of

Our study is most closely related to Zeckhauser and Fels (1968), Kimball (1987), and Bergstrom (1999), who analyzed the question whether systems of altruistically interdependent utility functions, in an intergenerational context, determine utilities as functions of allocations. In a similar vein, Hori (2001) investigated the same question in some more generality in the case of finitely many decision-makers. However, to the best of our knowledge, the results presented here are all new.

More exactly, we here consider a sequence of decision-makers, one in each time period. These decision-makers could be successive generations or the successive incarnations of one and the same individual. The decision-makers are assumed to have preferences that can be represented as the discounted sum of instantaneous utilities (for example from consumption). As in most macroeconomic models, the instantaneous utility function is the same in all periods, and all decision-makers use the same discount function over their respective futures. Let thus $f(t) \in [0, 1]$ be the *discount factor* that a decision-maker attaches to the instantaneous utility t periods ahead. Such representations of time preferences are commonplace in the economics literature. For example, in the seminal paper by Samuelson (1937), $f(t) = \delta^t$ for some $\delta \in (0, 1)$, while $f(t) = \beta\delta^t$ for some $\beta \in (0, 1]$ and $\delta \in (0, 1)$ in Phelps and Pollack (1968) and Laibson (1997). We will refer to the first case as *exponential discounting* and call the second, more general case, *quasi-exponential discounting*.⁴

We ask whether utility functions of this form are consistent with pure altruism towards future decision-makers in the sense that each decision-maker's utility can be written as the sum of own instantaneous utility (from, say, current consumption), and some weighted sum of all future decision-makers' total utilities. Let $a(t)$ be the *altruism weight* that the current decision-maker implicitly places on the total utility of his or her t :th successor.⁵ In Edgeworth's (1881) words, cited above, $a(t)$ is the current decision-maker's *coefficient of effective sympathy* for the decision-maker t periods later.

We find that such a function a always exists and that its values can be determined from a relatively simple recursive equation (proposition 1). However, there is no guarantee, *a priori*, that all function values are non-negative. A negative function value $a(t)$ means that the decision-maker is "spiteful" to his or her t :s successor, that is, the decision-maker

limited attention a systematic analysis of the relationship between these two frameworks is yet to be written, and appears to be quite a challenge, especially for models with an infinite horizon." (Ray, 1987, pp. 113-114)

⁴Such discount functions are frequently called hyperbolic or quasi-hyperbolic. We prefer the present terminology since this class of functions contain exponential (but not hyperbolic) functions as special cases.

⁵Altruistic concern for *earlier selves* seems irrelevant since earlier selves, by definition, do not exist at the time of the decision in question. The same holds for *earlier generations*, unless some ancestor is still alive when the current generation makes its decisions, as in overlapping generations models. Kimball (1987) analyses such a model.

Note, however, that if a decision-maker derives utility from memories of (his or his ancestors) past consumption, then even a strictly forward-looking altruistic decision-maker may rationally "invest" in future memories. However, that falls outside the scope of the present study - see Kimball (1987) and Ray and Wang (2002) for discussions of these issues.

prefers allocations that make that later decision-maker *worse* off.⁶ It is thus desirable to identify conditions on the discount function f that guarantee that all function values $a(t)$ be nonnegative. It turns out that a sufficient condition for this is that f be positive and that the ratio $g(t) = f(t)/f(t-1)$ between successive discount factors be non-decreasing in t (proposition 2). This ratio reflects the decision-maker's *patience* concerning events t periods ahead: it expresses the decision maker's dislike of a one-period postponement as a function of how many periods ahead this delay is to occur. The condition thus requires decision-makers to be more patient with postponements that are further away in the future - a property that seems to conform with all available empirical evidence (see for example Frederick et al. (2001)). Moreover, the condition is trivially met by exponential discounting - since then g is constant - and, more generally, by all quasi-exponential discount functions. Moreover, the conditions is also met by some, but not all, hyperbolic discount functions discussed in the psychology literature on time preferences (see for example Ainslie (1992)). We also show that our "patience" condition is closely related to, but distinct from, convexity of the discount function f (proposition 4).

As is well-known, exponential discounting - the canonical model in the economics literature - corresponds to one-period pure altruism: each decision-maker attaches a positive weight to his or her successor's utility and zero weight to all other decision-makers. We show that, in terms of pure altruism, exponential discounting is a boundary case in the following sense: a necessary condition for altruism towards future decision-makers is that future periods should not be more heavily discounted than what is obtained by exponential extrapolation of the discounting from the present to the next period (proposition 3). Also this necessary condition seems to agree with empirical evidence. However, the condition is evidently violated by certain discount functions used in the economics literature, namely those that place positive weight on instantaneous utility in some nearby period, but zero weight on instantaneous utility in some more distant period (as, for example, when each generation only cares about its own and the next generation's consumption).

We also find that quasi-exponential discounting in the Laibson-Phelps-Pollak (β, δ) -form corresponds to *exponential altruism*: the implied altruism weights decline exponentially over all future decision-makers. Hence, while the special case $\beta = 1$ of pure exponential discounting corresponds to altruism to the next generation only, $\beta < 1$ corresponds to (positive) exponential altruism to *all* future generations.

The remainder of the paper is organized as follows. The model is set up in section 2, and section 3 presents our results. Section 4 analyzes a few examples, and section 5 concludes. Mathematical proofs are provided in an appendix.

⁶Such "spite" may be justified in the context of addiction, however, where the current self may hold a paternalistic disregard for the preferences of the "hooked" future self.

2 Model

Consider a sequence of decision-makers $\tau = 0, 1, 2, \dots$. There may be finitely or infinitely many such decision-makers. However, in order to save on notation and treat the most challenging case, we henceforth presume an infinite sequence.⁷ Suppose, thus, that in each time period $t \in \mathbb{N} = \{0, 1, 2, \dots\}$ there is a single decision-maker who takes some action $x_t \in X$, where X is the set of alternatives available in each period t (for example, X may be a set of relevant consumption bundles). For the sake of concreteness, we will call action x_t *consumption* in period t , and by a *consumption stream* (or *allocation*) x we mean the infinite sequence of consumption vectors x_t , $x = (x_0, x_1, \dots) \in X^\infty$.

Each decision-maker τ has preferences \succsim_τ over consumption streams $x \in X^\infty$. A *preference profile* \succsim for the sequence of decision-makers is thus a sequence $\langle \succsim_\tau \rangle_{\tau \in \mathbb{N}}$ of preferences, one for each decision-maker τ . We here focus on preference profiles $\langle \succsim_\tau \rangle_{\tau \in \mathbb{N}}$ for which there exist functions $U_\tau : X^\infty \rightarrow \mathbb{R}$, one for each decision-maker τ , such that $x \succsim_\tau y$ if and only if $U_\tau(x) \geq U_\tau(y)$, where

$$U_\tau(x) = \sum_{t=0}^{\infty} f(t)u(x_{\tau+t}) \quad (1)$$

for some $u : X \rightarrow \mathbb{R}$ and $f : \mathbb{N} \rightarrow \mathbb{R}_+$ with $f(0) = 1$. We will call $u(x_s)$ the *instantaneous (sub)utility* from consumption in period s , and $f(t)$ the *discount factor* that each decision-maker attaches to the instantaneous utility t periods later. Hence, each decision-maker uses the same instantaneous subutility function u and discount function f .

We will say that a sequence $\langle U_\tau \rangle_{\tau \in \mathbb{N}}$ of utility functions (1) is *consistent with* (additively separable) *pure altruism* if for all $\tau \in \mathbb{N}$ and $x \in X^\infty$,

$$U_\tau(x) = u(x_\tau) + \sum_{t=1}^{\infty} a(t)U_{\tau+t}(x) , \quad (2)$$

for some $a : \mathbb{N}_+ \rightarrow \mathbb{R}_+$, where $\mathbb{N}_+ = \{1, 2, \dots\}$. Here $a(t)$ will be called the *altruism weight* that the decision-maker places on the welfare or total utility of the decision-maker t periods later.⁸

3 Results

Under what conditions is a sequence $\langle U_\tau \rangle_{\tau \in \mathbb{N}}$ of utility functions, defined in equation (1), consistent with pure altruism, and if it is, what are the implied altruism weights? A key result for answering this and related questions is the following observation:

⁷All results are easily adapted to the case of a finite number of decision makers.

⁸The restriction to additively separable altruism is not binding in the present context, see proposition 1.

Proposition 1 *If $\langle U_\tau \rangle$ is a sequence of real-valued utility functions satisfying equation (1) for some $u : X \rightarrow \mathbb{R}$ and $f : \mathbb{N} \rightarrow \mathbb{R}$, then $\langle U_\tau \rangle$ also satisfies equation (2), where $a : \mathbb{N}_+ \rightarrow \mathbb{R}$ is the unique solution to*

$$a(t) = \begin{cases} f(1) & \text{if } t = 1 \\ f(t) - \sum_{s=1}^{t-1} f(t-s)a(s) & \text{if } t > 1 \end{cases} . \quad (3)$$

Clearly, the discount function f may be recovered from the altruism function a from equation (3):

$$f(t) = \begin{cases} a(1) & \text{if } t = 1 \\ \sum_{s=0}^{t-1} a(t-s)f(s) & \text{if } t > 1 \end{cases} \quad (4)$$

This recursive equation, which determines f from a (recall the normalization $f(0) = 1$), essentially states that the discount factor $f(t)$ attached to the instantaneous utility t periods later equals that period's contribution to the utility of all interim decision-makers, weighted by their respective altruism weights. For example, the discount factor $f(2)$ equals the altruism weight placed on the decision-maker two periods ahead, plus the altruism weight placed on the decision-maker one period ahead times that decision-maker's one-period discounting: $f(2) = a(2) + a(1)f(1) = a(2) + a^2(1)$.

It is immediate from equation (4) that if the function a is nonnegative, so is f . However, as pointed out above, and seen in equation (3), a may well take negative values. For example, $a(2) = f(2) - a(1)f(1) = f(2) - f^2(1)$, so in order for $a(2)$ to be negative it suffices that $f(2) < f^2(1)$.⁹ In particular, this is the case when $f(1) > 0$ and $f(2) = 0$, as in models where each generation's welfare is a function of its own consumption and that of its immediate descendant. Another example of negative utility weights $a(t)$ is when $f(0) = 1$ and $f(t) = 1/(0.5 + t)$ for all $t > 0$; again $f^2(1) > f(2)$. A third example is $f(t) = 1/(1 + t^2)$ for all t . A fourth example is when the parameter β in the quasi-exponential (β, δ) -representation kicks in with one period's delay, that is, when $f(0) = 1$, $f(1) = \delta$ and $f(t) = \beta\delta^t$ for all $t \geq 2$. Then $a(2) = (\beta - 1)\delta^2 < 0$ for all $\beta < 1$.

For what class of discount functions f can one then guarantee that all welfare weights $a(t)$ are nonnegative? It turns out that a sufficient condition for this is that f be everywhere positive and that the associated "patience" function $g : \mathbb{N}_+ \rightarrow \mathbb{R}$, defined by $g(t) = f(t)/f(t-1)$, be non-decreasing. This condition is clearly met by all quasi-exponential discount functions with $\beta \leq 1$.

Proposition 2 *Suppose $f > 0$, and let g be the associated patience function. If g is non-decreasing, then $a \geq 0$. If g is strictly increasing, then $a > 0$.*

We note that if both f and a are non-negative - as under the hypothesis of proposition 2 - then the altruistic weight $a(t)$ attached to the decision-maker in any period t cannot exceed the discount factor $f(t)$ attached to consumption in that period; by equation (3)

⁹More generally, it is easy to verify that equation (3) gives $a(t) = (-1)^{t+1}f(1)^t$ for all $t \geq 1$ if $f(t) = 0$ for all $t > 1$.

we have $0 \leq a(t) \leq f(t)$ for all $t > 0$. In particular, if $f(t)$ goes to zero as t goes to infinity, then so does $a(t)$.

Another observation is that the hypothesis in proposition 2 is closely related to, but distinct from, the condition that f be convex, where we call a function $f : \mathbb{N} \rightarrow \mathbb{R}$ *convex* if its piece-wise affine extension to \mathbb{R}_+ is convex.¹⁰ It is easy to see that convexity is not sufficient for the altruism weights to be nonnegative. For example, any convex function with $f(0) = 1$, $f(1) = 0.2$ and $f(2) = 0.02$ has $a(2) = -0.02$. However, a seemingly slight strengthening of the hypothesis in proposition 2 effectively requires f to be convex. To see this, let $\tilde{f} : \mathbb{R}_+ \rightarrow \mathbb{R}$ be twice differentiable with $\tilde{f}(0) = 1$, $\tilde{f} > 0$ and $\tilde{f}' \leq 0$, and let $G : \mathbb{R}_+^2 \rightarrow \mathbb{R}$ be defined by

$$G(t, s) = \frac{\tilde{f}(t+s)}{\tilde{f}(t)}. \quad (5)$$

Clearly, the restriction of \tilde{f} to \mathbb{N} is a discount function f , and the associated patience function g satisfies $g(t+1) = G(t, 1)$ for all $t \in \mathbb{N}$. In particular, g is non-decreasing - as required in proposition 2 - if G'_1 , the partial derivative of G with respect to its first argument, is nonnegative. Under the latter, somewhat more stringent hypothesis, \tilde{f} , and hence also f , are convex (and thus, by proposition 2, a is nonnegative):

Proposition 3 *If $\tilde{f} : \mathbb{R}_+ \rightarrow \mathbb{R}$ is twice differentiable with $\tilde{f}(0) = 1$, $\tilde{f} > 0$, $\tilde{f}' \leq 0$, and $G'_1 \geq 0$, then f is convex and g non-decreasing.*

We next turn to the task of identifying a *necessary* condition for all altruism weights to be nonnegative. For this purpose, let us first briefly consider the classical case when the discount function f is exponential: $f(t) = \delta^t$ for all t , for some $\delta \in (0, 1)$. It is not difficult to verify by induction that equation (3) then gives $a(1) = \delta$ and $a(t) = 0$ for all integers $t > 1$. In other words, exponential discounting is equivalent to altruism to the next decision-maker only.

Conversely, one-period altruism implies exponential discounting: if $a(1) = \alpha \geq 0$ and $a(t) = 0$ for all $t > 1$, then $f(t) = \alpha^t$ for all t (see equation (4)). This is not surprising: if each decision-maker attaches an altruistic weight α to the next decision-maker, and zero weight to all others, then the contribution to current welfare from the instantaneous utility t periods ahead should be the product of how much the current decision-maker cares about the next decision-maker, how much the next decision-maker cares about his successor, and so on, up to the t 'th decision maker.

These observations concerning exponential discounting can be used to establish a necessary condition for (nonnegative) altruism in general, namely, that the discount function should not decline faster than exponentially, as compared with its decline from the current

¹⁰For all reals t between any two integers k and $k+1$, let

$$f^*(t) = (t-k)f(k+1) + (k+1-t)f(k),$$

thus defining a piece-wise affine extension f^* of f .

period to the next. Hence, exponential discounting is a boundary case from the viewpoint of altruism.

Proposition 4 *If $a \geq 0$, then $f(t) \geq [f(1)]^t$ for all t .*

4 Examples

4.1 Quasi-exponential discounting

What altruism weights correspond to the quasi-exponential discounting in the Laibson-Phelps-Pollak model? Suppose, thus, that $f(0) = 1$ and $f(t) = \beta\delta^t$ for all positive integers t , for some $\beta \in (0, 1]$ and $\delta \in (0, 1)$. Then $a(1) = \beta\delta$ and $a(2) = \beta(1 - \beta)\delta^2$. By induction in t , it is easily verified that:¹¹

$$a(t) = \beta(1 - \beta)^{t-1}\delta^t \quad \forall t \geq 1. \quad (6)$$

Not surprisingly, *one-period altruism* is obtained when $\beta = 1$: then $a(1) = \delta$ and $a(t) = 0$ for all $t > 1$. By contrast, when $\beta < 1$, then one obtains *exponential altruism*:

$$U_\tau(x) = u(x_\tau) + \alpha \sum_{t=1}^{\infty} \gamma^t U_{\tau+t}(x) , \quad (7)$$

where $\alpha = \beta/(1 - \beta)$ and $\gamma = (1 - \beta)\delta$. Here γ is the constant factor by which altruism declines over the infinite sequence of future decision-makers. The total utility weight that (7) places on the aggregate of all future decision-makers is $A = \alpha\gamma/(1 - \gamma)$. We note that $\gamma \rightarrow 0$ and $\alpha\gamma \rightarrow \delta$ as $\beta \rightarrow 1$. In other words, the representation is continuous at $\beta = 1$ - the boundary case of classical exponential discounting.

Conversely, if the sequence $\langle U_\tau \rangle$ of utility functions satisfy (7) for some $\alpha > 0$ and $0 < \gamma < 1$, then the underlying preferences are behaviorally equivalent with quasi-exponential discounting with $\beta = \alpha/(\alpha + 1)$ and $\delta = (\alpha + 1)\gamma$. In particular, the induced discount function f is summable iff $\gamma < 1/(\alpha + 1)$, that is, iff the altruism weights do not taper off too slowly, given α .¹² We also note that if decision makers place altruism weight $1/2^t$ on its t :th successor - that is, the genetic kinship factor between parent and child - then $\alpha = 1$, $\gamma = 1/2$, $A = 1$, $\beta = 1/2$ and $\delta = 1$. Hence, in this border-line case, f is not

¹¹Suppose $a(s) = \beta(1 - \beta)^{s-1}\delta^s$ for $s \leq t$. Then (4) gives

$$a(t+1) = \beta\delta^{t+1} - \beta^2\delta^{t+1} \sum_{s=1}^t (1 - \beta)^{s-1},$$

which, after simplification, boils down to $a(t+1) = \beta(1 - \beta)^t\delta^{t+1}$.

¹²A function $f : \mathbb{N} \rightarrow \mathbb{R}$ is *summable* if $\sum_{t=0}^T f(t)$ converges to some real number as $T \rightarrow \infty$. Summability is necessary in order for the utility functions U_τ to be defined for constant consumption streams.

summable - the utility functions U_τ are then not well-defined for constant consumption streams.¹³

For the sake of illustration of this one-to-one relationship between discount factors and altruism weights, we note that Angeletos et al (2001) made the following estimate of the parameter pair (β, δ) in the Laibson-Phelps-Pollak model, based on annual US data: $\beta = 0.55$ and $\delta = 0.96$. The associated altruism parameters (for annual decision makers) are thus $\alpha \approx 1.22$ and $\gamma \approx 0.43$, which yields total altruism $A \approx 0.92$ towards the aggregate of future decision makers.

4.2 Hyperbolic discounting

Psychologists who have studied temporal preferences of human and animal subjects suggest that the discount function f be hyperbolic, see Herrnstein (1981), Mazur (1987) and Ainslie (1992). In this vein, Loewenstein and Prelec (1992) suggested discount functions f of the form $f(t) = (1 + \mu t)^{-\gamma}$ for $\mu, \gamma > 0$. It is easily verified that the associated patience function g is increasing. Hence, all discount functions of the Loewenstein and Prelec variety are consistent with altruism towards future decision-makers.¹⁴ By contrast, some other, closely related, discount functions are not, as was noted in section 3 above. There, we saw that if $f(0) = 1$ and $f(t) = (\lambda + \mu t)^{-\gamma}$ for all $t > 0$, then the monotonicity condition for g is violated when $\lambda = 0.5$ and $\mu = \gamma = 1$. (Note that f is non-increasing if $\lambda, \mu \geq 0$ and $\lambda + \mu \geq 1$, which we now assume.) In this class of discount functions, the sufficient condition in proposition 2 is in fact also necessary: if $(\lambda + \mu)^2 < \lambda + 2\mu$, then $a(2) < 0$.¹⁵ That proposition is thus sharp within this class of discount functions. Actually, proposition 4 is sharp as well: the necessary condition for (nonnegative) altruism in that proposition is equivalent with the condition that $(\lambda + \mu)^t \geq \lambda + \mu t$ for all $t > 0$, which, in its turn, is equivalent with $(\lambda + \mu)^2 \geq \lambda + 2\mu$.¹⁶

5 Conclusions

We started out by asking if discounting of future instantaneous utilities is consistent with “pure” altruism towards future decision-makers. We identified a recursive functional equation which establishes a one-to-one relationship between discount factors and altruistic weights attached to future generations or future selves. We saw that some discount functions used in the literature are consistent with altruism towards one’s future selves or future generations, while others are not. We also established a sufficient condition, and

¹³Note that summability is irrelevant if the intertemporal budget constraint requires finite total consumption of all goods ($\sum_t x_{ti} < +\infty$ for all i), such as in the case of non-renewable resources.

¹⁴We have been unable to obtain a closed-form representation of the altruistic weights corresponding to hyperbolic discounting.

¹⁵This follows from (3), which gives $a(2) = f(2) - [f(1)]^2$, a difference which is nonnegative precisely when g is non-decreasing.

¹⁶To see this, note that the derivative of $(\lambda + \mu)^t - \lambda - \mu t$ with respect to t is nonnegative for all $t \geq 2$.

a necessary condition, for consistency in this respect. These conditions are met by the quasi-exponential discounting models currently under investigation in the macroeconomics literature (see, for example, Laibson (1997), Barro (1999), Laibson and Harris (2001) and Angeletos et al (2001)), as well as by some of the hyperbolic discounting models in the psychology literature (see, for example, Loewenstein and Prelec (1992)).

From a behavioral viewpoint, however, separable discounting models seem quite restrictive as representations of intertemporal consumer preferences. See, for example, Frederick, Loewenstein and O'Donoghue (2001), Kahneman (2000) and Rubinstein (2001) for alternative approaches to intertemporal choice. We hope, however, that our study has shed some light on a somewhat wider path than the well-trodden but narrow path of exponential discounting.

6 Appendix

6.1 Proof of proposition 1

Suppose $\langle U_\tau \rangle$ satisfies equation (1) for some $u : X \rightarrow \mathbb{R}$ and $f : \mathbb{N} \rightarrow \mathbb{R}$ with $f(0) = 1$. Let $a : \mathbb{N}_+ \rightarrow \mathbb{R}$ be defined by (3). Then

$$f(t) = \sum_{s=1}^t a(s)f(t-s) \quad \forall t \in \mathbb{N}_+ . \quad (8)$$

Hence,

$$\begin{aligned} U_\tau(x) &= u(x_\tau) + \sum_{t=1}^{\infty} \sum_{s=1}^t a(s)f(t-s)u(x_{\tau+t}) = \\ &= u(x_\tau) + \sum_{s=1}^{\infty} a(s) \left[\sum_{t=s}^{\infty} f(t-s)u(x_{\tau+t}) \right] \\ &= u(x_\tau) + \sum_{s=1}^{\infty} a(s) \left[\sum_{k=0}^{\infty} f(k)u(x_{\tau+s+k}) \right] = u(x_\tau) + \sum_{s=1}^{\infty} a(s)U_{\tau+s}(x) \end{aligned} \quad (9)$$

Since the resulting equation holds for all τ , this proves the claim.¹⁷

6.2 Proof of proposition 2

Suppose first that g is non-decreasing. Since $a(1) = f(1)$, we have $a(1) > 0$. Suppose $a(s) \geq 0 \quad \forall s < t$. By equation (3):

¹⁷It is easily verified that the change of order of summation in this derivation is justified. For by assumption U_τ is a real-valued function and hence the sum in the first line converges to some $\lambda \in \mathbb{R}$. Hence, for every $\varepsilon > 0$ there exists a T_ε such that summation from $t = 1$ up to T_ε brings this partial sum within ε from λ . Given T_ε , the order of summation may be changed and the final expression is obtained by letting $\varepsilon \rightarrow 0$.

$$\begin{aligned}
a(t) &= f(t) - f(1)a(t-1) - \sum_{s=1}^{t-2} a(s)f(t-s) \\
&= g(t)f(t-1) - f(1)a(t-1) - \sum_{s=1}^{t-2} g(t-s)a(s)f(t-s-1) \\
&\geq g(t) \left[f(t-1) - \sum_{s=1}^{t-2} a(s)f(t-s-1) \right] - f(1)a(t-1) \\
&= g(t)a(t-1) - f(1)a(t-1) \\
&= [g(t) - f(1)] a(t-1) \geq 0,
\end{aligned} \tag{10}$$

where equation (3) was used again for $a(t-1)$, and where the last inequality follows from the fact that, by assumption, g is non-decreasing with $g(1) = f(1)$.

Secondly, suppose that g is strictly increasing. If $a(s) > 0 \forall s \leq t$, then the same reasoning as above leads to $a(t) > [g(t) - g(1)] a(t-1) > 0$.

6.3 Proof of proposition 3

By differentiation,

$$G'_1(t, s) \geq 0 \quad \Leftrightarrow \quad \tilde{f}'(t+s)\tilde{f}(t) \geq \tilde{f}(t+s)\tilde{f}'(t). \tag{11}$$

Suppose that $G'_1(t, s) \geq 0$ for all $s > 0$. If $\tilde{f}'(t) = 0$, then $\tilde{f}'(t+s) \geq 0$ for all $s > 0$, and hence $\tilde{f}''(t) \geq 0$. If instead $\tilde{f}'(t) < 0$, then

$$\begin{aligned}
\tilde{f}''(t) &= \lim_{s \downarrow 0} \frac{\tilde{f}'(t+s) - \tilde{f}'(t)}{s} \geq \lim_{s \downarrow 0} \frac{1}{s} \left[\frac{\tilde{f}(t+s)}{\tilde{f}(t)} - 1 \right] \tilde{f}'(t) \\
&= \lim_{s \downarrow 0} \left[\frac{\tilde{f}(t+s) - \tilde{f}(t)}{s} \right] \frac{\tilde{f}'(t)}{\tilde{f}(t)} = \frac{[\tilde{f}'(t)]^2}{\tilde{f}(t)} > 0.
\end{aligned} \tag{12}$$

6.4 Proof of proposition 4

Suppose $a \geq 0$, and let f be the associated discount function, as defined in (4). Let $\alpha = a(1)$, and let a^* be the altruism-weight function defined by $a^*(1) = \alpha$ and $a^*(t) = 0$ for all $t > 1$. We know from the above observation that the discount function f^* associated with a^* is $f^*(t) = \alpha^t$ for all t . However, it follows from (4) that $f(t) \geq f^*(t)$ for all $t > 1$ since $a \geq a^*$. Hence, $f(t) \geq \alpha^t = [f(1)]^t$ for all $t > 1$.

References

- [1] Ainslie G. (1992): *Picoeconomics*. Cambridge University Press (Cambridge).
- [2] Andreoni (1989): “Giving with impure altruism: applications to charity and Ricardian equivalence”, *Journal of Political Economy* 97, 1447-1458.
- [3] Angeletos G.-M., D. Laibson, A. Repetto, J. Tobacman and S. Weinberg (2001): “The hyperbolic consumption model: calibration, simulation and empirical evaluation”, forthcoming, *Journal of Economic Perspectives*.
- [4] Barro R. (1974): “Are government bonds net wealth?”, *Journal of Political Economy* 82, 1095-1117.
- [5] Bergstrom T. (1999): “Systems of benevolent utility functions”, *Journal of Public Economic Theory* 1, 71-100.
- [6] Bernheim D. (1986): “On the voluntary and involuntary provision of public goods”, *American Economic Review* 76, 789-793.
- [7] Bernheim D. (1987): “Ricardian equivalence: An evaluation of theory and evidence”, in S. Fischer (ed.), *NBER Macroeconomics Annual* 2, MIT Press (Cambridge, USA).
- [8] Collard D. (1975): “Edgeworth’s proposition on altruism”, *Economic Journal* 85, 355-360.
- [9] Edgeworth F. (1881): *Mathematical Psychics*. Kegan Paul (London).
- [10] Frederick S., G. Loewenstein and T. O’Donoghue (2002): “Time Discounting and Time Preference: A Critical Review”, *Journal of Economic Literature* 40, 351-401.
- [11] Herrnstein R. (1981): “Self control and response strength”, in C. Bradshaw and C. Lowe (eds.), *Quantification of Steady-State Operant Behavior*, Elsevier (Amsterdam).
- [12] Hori H. (2001). “Non-paternalistic altruism and utility interdependence”, *Japanese Economic Review* 52, 137-155.
- [13] Kahneman D. (2000): “Experience utility and objective happiness: a moment-based approach”, in D. Kahneman and A. Tversky, *Choices, Values and Frames*. Cambridge University Press (Cambridge).
- [14] Kimball M. (1987): “Making sense of two-sided altruism”, *Journal of Monetary Economics* 20, 301-326.
- [15] Laibson D. (1997): “Golden eggs and hyperbolic discounting”, *Quarterly Journal of Economics* 112, 443-447.

- [16] Laibson D. and C. Harris (2001): “Dynamic choices of hyperbolic consumers”, *Econometrica* 69, 935-957.
- [17] Loewenstein G. and D. Prelec (1992): “Anomalies in intertemporal choice: evidence and an interpretation”, *Quarterly Journal of Economics* 57, 573-598.
- [18] Mazur J. (1987): “An adjustment procedure for studying delayed reinforcement”, in J. Commons, J. Mazur and H. Rachlin, *The Effect of Delay and Intervening Events on Reinforcement Values*, Erlbaum (Hillsdale, NJ).
- [19] Phelps E. and R. Pollak (1968): “On second-best national savings and game-equilibrium growth”, *Review of Economic Studies* 35, 201-208.
- [20] Ray D. (1987): “Nonpaternalistic intergenerational altruism”, *Journal of Economic Theory* 41, 112-132.
- [21] Ray D and R Wang (2002), “On Some Implications of Backward Discounting”, mimeo, New York University.
- [22] Rubinstein A. (2001): “A theorist’s view of experiments”, *European Economic Review* 45, 615-628.
- [23] Samuelson P. (1937): “A note on measurement of utility”, *Review of Economic Studies* 4, 155-161.
- [24] Zeckhauser R. and S. Fels (1968): “Discounting for proximity with perfect and total altruism”, Harvard Institute of Economic Research, Discussion Paper No. 50.