CMS-EMS
Center for Mathematical Studies in Economics
And Management Science

Discussion Paper #1571

The Logical Consistency of Time Inconsistency
A Theory of Forward-Looking Behavior

Simone Galperti*
Bruno Strulovici**

October 29, 2013

JEL Classification: D01, D60, D90.

Keywords: time inconsistency, forward-looking behavior, hyperbolic discounting, beta-delta discounting, anticipations, welfare criterion.

* UC, San Diego
**Northwestern University

The Logical Consistency of Time Inconsistency

A Theory of Forward-Looking Behavior

Bruno Strulovici
Northwestern University

Simone Galperti*
UC, San Diego

First Draft: May 2011

October 29, 2013

Abstract

This paper argues that, to be forward-looking in a logically consistent sense, a decision maker must take account of his overall well-being, not just his instantaneous utility, in all future periods. However, such a decision-maker is necessarily time inconsistent. The paper explores the relationship between how a decision-maker discounts well-being and how he discounts instantaneous utility. It also provides simple axiomatizations of preferences that exhibit forward-looking behavior, including quasi-hyperbolic discounting (Phelps and Pollack (1968) [18]; Laibson (1997) [12]). Finally, the paper provides a rigorous way to think about welfare criteria in models with time inconsistent agents.

KEYWORDS: time inconsistency, forward-looking behavior, hyperbolic discounting, beta-delta discounting, anticipations, welfare criterion.

JEL CLASSIFICATION: D01, D60, D90.

*The authors are grateful to Nabil Al-Najjar, Eddie Dekel, Bart Lipman, Peter Klibanoff, Mark Machina, David G. Pearce, Todd Sarver, Marciano Siniscalchi, and Asher Wolinsky for useful comments.

1 Introduction

The standard theory of intertemporal choice hinges on the concept of instantaneous utility: the utility that a consumption event yields at the time when it occurs. According to the usual interpretation of this theory, at each time, a forward-looking decision maker (DM) assigns to any consumption stream an overall utility that aggregates instantaneous utilities from then on. Such a DM, however, is forward-looking in a very limited sense: He does not realize that, at any future time, his overall utility will continue to depend on immediate as well as future consumption.[1] This logical inconsistency calls for reconsidering how we think about and model forward-looking behavior. This paper aims to do so, by proposing a theory based on the broader concept of well-being. In this theory, a DM will be fully forward-looking, for at each time his overall utility (hereafter, 'well-being') depends not on his instantaneous utilities, but on his well-being at all future dates.[2]

To illustrate the conceptual issues raised by the standard theory, consider the geometrically discounted utility (DU) model (Samuelson (1937) [21]). To be concrete, imagine you have just won a trip to Hawaii, which will take place in two weeks. How does that affect your well-being today (time 0)? For simplicity, ignore other consumption and normalize the utility of consuming nothing to zero. The DU model says that, at time 0, your well-being is

$$U_0 = u(c_0) + \delta u(c_1) + \delta^2 u(c_2) + \cdots = \delta^2 u(c_2),$$

where δ is your discount rate, $u(\cdot)$ measures your instantaneous utility, and c_2 is the 'trip consumption.'[3] Yet, it also says that at time 1—i.e., one week before the trip—your well-being is

$$U_1 = u(c_1) + \delta u(c_2) + \cdots = \delta u(c_2),$$

which is not zero. So the DU model must imply that, at time 1, your well-being takes account of the Hawaii trip. Is this dependence recognized at time 0?

The answer depends on how we interpret the DU model. If we view the time-0 self as thinking about the time-1 self only in terms of time-1 consumption, which is zero, then

[1] An alternative interpretation will be discussed later. As will be clear, that interpretation is also problematic.

[2] Since the goal here is to address the logically inconsistent way in which the standard theory models forward-looking behavior, the paper focuses of the case in which current well-being depends on future well-being. Of course, it is natural to consider the case in which current well-being may also depend on past well-being. But this is beyond the scope of this paper.

[3] Allowing for time-dependent utility, u_t, is irrelevant for the main point here.

the answer is no. This interpretation is logically inconsistent, for it treats the time-1 self as forward-looking if we look at him from time 1, but as myopic if we look at him from time 0. Another interpretation is that time-0 well-being depends on time-0 consumption and time-1 well-being, i.e.,

$$U_0 = u(c_0) + \delta U_1.$$

In words, at time 0, you care about the trip *only* through the well-being that your time-1 self will get from thinking about going to Hawaii a week later. Accordingly, the time-0 self does realize that the time-1 self's well-being depends on the Hawaii trip, but he does not care about the trip per se. This interpretation, though logically consistent, is unsatisfactory. There is no reason why the time-0 self should not care directly about the time-2 self—or even later selves, for that matter.

To overcome these issues, this paper axiomatizes and analyzes preference representations that properly capture fully forward-looking behavior. In general, preferences over consumption streams, c, will be represented by the function

$$U(c) = V(c_0, U({}_1c), U({}_2c), \ldots), \tag{1}$$

where ${}_tc$ is the continuation of c from t onwards. This expression highlights the two conceptual premises of this paper. First, future well-being, not future instantaneous utility, determines today's well-being. Therefore, for instance, DM is insensitive to reshuffling consumption across dates, as long as this does not change future well-being. Of course, today's well-being also depends on today's consumption. The second premise is that immediate consumption is essentially different from (future) well-being, a more general and comprehensive entity. Such a distinction links this paper with some early work on intertemporal choice by Böhm-Bawerk (1890)[24] and Fisher (1930) [5], who had a similar point of view.

The present paper focuses on preferences that do not change over time. Nonetheless, fully forward-looking behavior turns out to be incompatible with time consistency. A sequence of preferences is time consistent, if it has the following property: If a course of action is preferable according to tomorrow's preference, then it remains preferable, for tomorrow, according to today's preference.[4] In this paper, this property holds if and only if V in (1) depends only on its first two arguments, so that DM does not care about his well-being beyond the immediate future. To see the intuition, consider

[4]Time consistency is conceptually different from stationarity of a preference (Koopmans (1960) [9]), even though these properties are often viewed as synonyms. This difference is discussed at the end of Section 4.

a setting with three periods, 1, 2, and 3.[5] In period 2, when choosing consumption for the last two periods, DM understands that his immediate well-being depends on period-3 well-being and acts accordingly. Instead, in period 1, when again considering how much to consume in the last two periods, DM understands that his immediate well-being depends on period-3 well-being in two ways: directly, but also indirectly via his period-2 well-being. So DM's view on consumption allocations in period 2 and 3 changes between period 1 and 2. Roughly speaking, in period 2, DM is unable to internalize the 'externality' that his choices have on himself in period 1—this period is, by then, part of the past.

This point implies that a fully forward-looking DM is necessarily time inconsistent. This conclusion may seem paradoxical, since such a DM cares *more* about his future well-being than in standard models, looking *beyond* the immediate future. This, perhaps unexpected, source of time inconsistency differs from usual rationales that view DM as overweighing the present vs. the future (Laibson (1997) [12]), or imply that he is prone to temptations (Gul and Pesendorfer (2001) [7]), or assume that his preferences change over time (Strotz (1955) [23]).

The paper then considers well-being representations that feature time discounting. The goal is to understand the implications of how DM discounts well-being on how he discounts instantaneous utility, and vice versa. As is well-known, geometric discounting of instantaneous utility implies time consistency; this, as noted, is equivalent to DM's being 'myopically' forward-looking. Therefore, this paper shows that we can interpret and explain departures from geometric discounting of instantaneous utility as situations in which DM is fully forward-looking. In this view, if we believe that usually people are fully forward-looking, then we should be less surprised by the overwhelming evidence against geometric discounting (see, e.g., Frederick et al. (2002) [6]).

The paper derives general formulae that allow us to express well-being discount functions only in terms of instantaneous-utility discount functions, and vice versa. In general, a fully forward-looking DM always discounts well-being more than he appears to discount instantaneous utility. The reason is simple: DM weighs his instantaneous utility at some future date t more than his well-being at t, because consumption at t affects well-being at t as well as at all previous periods.

To illustrate how these formulae work, the paper considers two renown cases: hyperbolic (see, e.g., Loewenstein and Prelec (1992) [15]) and quasi-hyperbolic discounting

[5]The point holds with finitely as well as infinitely many periods. Of course, in the first case, preferences cannot be time invariant because their domain changes over time. So this example is just for illustration.

(Phelps and Pollack (1968) [18]; Laibson (1997) [12]). In the first case, a faster decline of the instantaneous-utility discount rate corresponds to slower discounting of well-being. In the second case, β-δ discounting corresponds to a specific well-being representation:[6]

$$U(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t \gamma U({}_t c), \tag{2}$$

where $\gamma = \frac{\beta}{1-\beta}$. Most importantly, the paper derives this representation from straightforward axioms on preferences. It thus provides a surprising, yet natural, axiomatization of quasi-hyperbolic discounting, which complements existing ones (see Section 2). The degree of present bias β is tightly linked with the parameter γ, interpreted as the 'vividness' or 'imaginability' of future well-being.[7] This link implies, for instance, that increasing the vividness of the future well-being caused by some behavior may mitigate present bias. This differs from providing information on the consequences of some behavior and could be done with appropriate ad campaigns, for issues like smoking, dieting, or even climate change.

Built on the concept of well-being, the theory in this paper can accommodate phenomena that appear as anomalies in the DU model.[8] For instance, despite discounting his future well-being, a fully forward-looking DM may desire to advance dreadful events (like a colonoscopy) and to postpone delightful events (like a party with friends). More generally, such a DM may prefer consumption streams that induce increasing, rather than decreasing, instantaneous-utility streams, while the DU model would predict the opposite. This is because later consumption improves well-being in more periods and therefore is more valuable from an ex-ante point of view.

Finally, this paper suggests a different angle from which we can think about welfare criteria, in contexts with time-inconsistent preferences. If a preference has a well-being representation, then we know explicitly to what extent DM's choices today take account of future well-being. In contrast, this information is not conveyed by a representation that depends only on instantaneous utility. This information, however, may be valuable to assess whether today's choices serve as a reasonable welfare criterion. For example, by expression (2), today's choices of a β-δ DM actually take account of his well-being in all future periods—not just tomorrow. As a result, one may argue that the β-δ model defines a more reasonable welfare criterion than the DU model.

[6] For a similar result, see Saez-Marti (2005) [20].

[7] This interpretation is related to Loewenstein (1987) [14] (see Section 2).

[8] For evidence of these phenomena, see, e.g., Frederick et al. (2002) [6].

As noted, the paper provides straightforward axiomatizations of several well-being representations. The general representation in (1) hinges on one simple axiom, which intuitively says the following. Given two consumption streams with equal initial consumption, if DM is indifferent between their continuations ($_tc$) starting at any future date, then he cannot prefer either stream—for example, because it allocates future consumption differently over time. This ensures that current well-being depends on future consumption only through the implied future well-being.

The paper then derives well-being representations characterized by time separability, discounting, and stationary dependence on future well-being. Time separability in immediate consumption and future well-being follows from a straightforward adaptation of Debreu's (1960) [3] strong-separability axiom. With regard to stationarity, intuitively, it says that, from the ex-ante point of view, DM does not see himself as changing how he ranks his options, simply because all options are postponed by one period. Of course, for this property to be reasonable, the time shift should not change the nature of the considered options. More concretely, from the point of view of this paper, postponing consumption streams that start at some future date does not change their nature, for it involves shifting future well-being. Instead, postponing consumption streams that start today changes their nature, for it transforms immediate consumption into future well-being. For this reason, in this paper, the stationarity axiom refers only to future well-being. This axiom differs substantively from Koopmans' (1960) stationarity, as well as from Olea and Strzalecki's (2013) quasistationarity (see Section 2).

Finally, to obtain the 'vividness' representation in (2) a further axiom is needed. Even if DM's preference is separable in immediate consumption and future well-being, the consumption trade-offs between two periods can be influenced by well-being—hence consumption—in future periods. This interdependence, although consistent with some evidence (see Frederick et al. (2002) [6]), is clearly absent in the β-δ model. So the additional axiom simply rules it out. The resulting axiomatic foundation of the β-δ model sheds light on its main features. First, the stark, apparently ad-hoc, distinction between short and long-run discounting of instantaneous utility is explained by the conceptual distinction between immediate consumption and future well-being, combined with stationarity in future well-being. Second, the time separability in instantaneous utility is explained by separability in immediate consumption and future well-being, combined with the impossibility that intertemporal consumption trade-offs depend on future well-being.

2 Related Literature

Several papers have addressed, in various ways, the criticisms that empirical studies have raised against the standard DU model. More specifically, Loewenstein (1987) [14], Caplin and Leahy (2001) [1], and Koszegi (2010) [11] augment the DU model with a psychological dimension called anticipation. In these papers, consumption at date t generates two kinds of utilities, both geometrically discounted: standard instantaneous utility at t and anticipation utility at all $t' < t$. Anticipation utility captures the effects on DM's happiness of feelings, like excitement or anxiety, that he harbors when thinking about future events. Such a psychological connotation sets anticipation apart from forward-looking behavior as intended here: Being forward-looking simply means that DM cares about the future per se—as in the DU model—and correctly forecasts that he continues to do so in the future. Although conceptually very different, anticipation models can also lead to time inconsistency and can accommodate the anomalies of the DU model addressed in the present paper. This paper, however, does so without augmenting the model with psychological dimensions. It therefore complements the explanations offered by anticipation models.

The result that fully forward-looking behavior leads to time inconsistency of preferences relates to the work on nonpaternalistic sympathy of Pearce (1983) [17]. This paper considers a cake-eating model with finitely many generations; each generation's well-being depends on its consumption as well as the well-being of all other or only future generations. The paper shows that any consumption plan in a non-cooperative equilibrium among generations must be inefficient. It leaves open, however, the question of whether equilibria are inefficient because the generations cannot coordinate on a better outcome—intuitively, like in the prisoner's dilemma—or because they pursue fundamentally inconsistent goals. The present paper shows that the second reason is the answer.

Phelps and Pollack (1968) [18] introduced β-δ discounting, so as to analyze economies populated by non-overlapping generations that are 'imperfectly altruistic.' That is, this paper views as more plausible that the current generation cares significantly more about its consumption than about that of any future generation—composed, after all, of unborn strangers. In this view, the β-δ formula, though simple, has a natural justification. Laibson (1997) [12] makes a significant conceptual leap, by applying the same formula to individual decision-making. He justifies using it based on its ability "to capture the qualitative properties" of hyperbolic discounting, which has received substantial empirical support. Nonetheless, with a single DM, it is more difficult to justify why he cares significantly more about his immediate consumption than *his own* future consumption in a uniform way.

Several papers have provided axiomatic justifications for the β-δ model in settings with a single DM (see, e.g., Hayashi (2003) [8]; Olea and Strzalecki (2013) [16]). These axiomatizations differ from that of the present paper as follows. First, they continue to view DM as caring about instantaneous utilities. Within this framework, they replace Koopmans' (1960) stationarity—clearly violated by β-δ discounting—with quasistationarity, namely stationarity from the second period onward. Assuming quasistationarity, however, does not address the conceptual issue posed by Laibson's model. Second, to obtain the β-δ representation, they need to ensure that current and future instantaneous utilities are cardinally equivalent. Olea and Strzalecki's clever axioms turn out to suggest simple experiments to identify and measure β and δ, but they are difficult to interpret. In contrast, the present paper views DM as caring about immediate consumption and future well-being. As a result, its notion of stationarity—involving only well-being—has a more natural interpretation and, as noted, can offer a rationale for the β-δ formula in Laibson's setting. The paper also shows that this formula is tightly linked with an intuitive property, namely that intertemporal consumption trade-offs do not depend on future well-being.

Finally, as noted, the literature has already discovered the mathematical equivalence between the β-δ formula and the representation in (2). For example, Saez-Marti and Weibull (2005) [20] derive this equivalence, while studying when discounting instantaneous utility is consistent with altruism, "in the sense that each DM's utility can be written as the sum of own instantaneous utility, and some weighted sum of all future DMs' total utilities." However, note that, in their paper, the weights $\alpha^t\gamma$ represent the degree of altruism of generation 0 towards generation t; so their interpretation is completely different from that in the present paper. More generally, Saez-Marti and Weibull do not provide an axiomatic foundation of either representation; hence, they also do not address the conceptual issues mentioned before.

3 Preliminaries

This paper considers how a decision maker (DM) evaluates consumption streams. At each time t, the set of feasible consumption levels is X. Time is discrete with infinite horizon: $T = \{0, 1, 2, \ldots\}$. The set of consumption streams is $C = X^T$.[9] For $t \geq 0$,

[9] The mathematical properties of X and C will be introduced later in Section 7.

let ${}_tC$ be the set of consumption streams starting at time t—elements of this set will be denoted by ${}_tc = (c_t, c_{t+1}, \ldots)$. At each t, DM has a preference relation $\succ^t$, defined over the domain ${}_tC$. This preference represents DM's choices over consumption streams at t. In principle, DM's preferences could very well be different at different times. This paper assumes, however, that they are time invariant.

Assumption 1 (Time Invariance)**.** *For all* $t \geq 0$, $\succ^t = \succ$.

This paper considers a DM who, at each t, cares about his future well-being (i.e., overall utility). That is, it focuses on preferences $\succ$ that have the following well-being representation. This, as well as other representations to follow, will be axiomatized in Section 7.

Definition 1 (Well-Being Representation)**.** The preference $\succ$ has a well-being representation if and only if there exist (continuous) functions $U : C \to \mathbb{R}$, with range $\mathcal{U}$, and $V : X \times \mathcal{U}^{\mathbb{N}} \to \mathbb{R}$, such that for any t

$$ {}_tc \succ {}_tc' \iff V(c_t, U({}_{t+1}c), U({}_{t+2}c), \ldots) > V(c'_t, U({}_{t+1}c'), U({}_{t+2}c'), \ldots), $$

and $U({}_tc) = V(c_t, U({}_{t+1}c), U({}_{t+2}c), \ldots)$.[10]

So the utility DM assigns to a consumption stream ${}_tc$ depends *only* on his immediate consumption c_t and, for at least some periods after t, on the per-period well-being that he will get from ${}_tc$. In this paper, well-being captures an aggregate of sensorial pleasure from immediate consumption and purely mental satisfaction (or dissatisfaction) from future well-being. While such an aggregation seems difficult to make, it is somehow performed by any forward-looking DM who must choose current consumption in a dynamic optimization problem. The definition is recursive, as one should expect: For a forward-looking DM, well-being today involves well-being in the future. Finally, note that this paper does not consider a DM who is fully myopic and cares only about immediate consumption, and that, at this stage, V may be strictly decreasing in $U({}_{t+k}c)$ for some $k \geq 1$.

4 Forward-Looking Behavior and Time (In)consistency

Among the properties of dynamic choice, time consistency is perhaps the most prominent and studied one. Therefore, this paper first considers what implications time consistency

[10] Of course, one can allow DM's preferences to differ across time, yet obtain for each $\succ^t$ a well-being representation. In this case, U and V will depend on t.

has on the extent to which DM is forward-looking—that is, on how V depends on future well-being. Surprisingly, there is a tension between being time consistent and being forward-looking.

In the present setup, the idea that choices are consistent over time is captured as follows (see, e.g., Siniscalchi (2011) [22]).

Definition 2 (Time Consistency)**.** If $({}_{t+1}c) \sim^{t+1} ({}_{t+1}c')$, then $(c_t, {}_{t+1}c) \sim^t (c_t, {}_{t+1}c')$. If $({}_{t+1}c) \succ^{t+1} ({}_{t+1}c')$, then $(c_t, {}_{t+1}c) \succ^t (c_t, {}_{t+1}c')$.[11]

This condition is based on the following question. Suppose DM is considering, at t, the consumption streams he will be facing from $t+1$ onward. Does DM rank such streams in the same way at t and at $t+1$? Importantly, these rankings are those revealed by DM's choice correspondences at t and $t+1$—which, by Assumption 1, are generated by the same preference. So Definition 2 first requires that, if at $t+1$ DM chooses either ${}_{t+1}c$ or ${}_{t+1}c'$, then at t he also choose either stream extended with the same consumption at t. In other words, a strict preference between $(c_t, {}_{t+1}c)$ and $(c_t, {}_{t+1}c')$ would be equivalent to different choice correspondences at t and at $t+1$, which is at odds with time consistency. On the other hand, if DM chooses ${}_{t+1}c$ but not ${}_{t+1}c'$ at $t+1$, then it is again a violation of time consistency if he chooses $(c_t, {}_{t+1}c')$.[12]

Time consistency has strong implications on the degree of forward-looking behavior. Indeed, DM's preference is time consistent if and only if he cares about immediate consumption and (positively) about his well-being in the next period only.

Proposition 1. *The preference $\succ$ satisfies time consistency if and only if $V(c_t, U({}_{t+1}c), U({}_{t+2}c), \ldots) = V(c_t, U({}_{t+1}c))$, for all $c \in C$, and V is strictly increasing in its second argument.*

Proof. Suppose $V(c_t, U({}_{t+1}c), U({}_{t+2}c), \ldots) = V(c_t, U({}_{t+1}c))$ for all $c \in C$ and V is increasing in the second argument. Using Assumption 1, if ${}_{t+1}c \sim {}_{t+1}c'$, then[13] $U({}_{t+1}c) = U({}_{t+1}c')$ and, since V is a function, $V(c_t, U({}_{t+1}c)) = V(c_t, U({}_{t+1}c'))$. This implies $(c_t, {}_{t+1}c) \sim (c_t, {}_{t+1}c')$. If ${}_{t+1}c \succ {}_{t+1}c'$, then $U({}_{t+1}c) > U({}_{t+1}c')$ and, since V is strictly increasing in its second argument, $V(c_t, U({}_{t+1}c)) > V(c_t, U({}_{t+1}c'))$; this implies $(c_t, {}_{t+1}c) \succ (c_t, {}_{t+1}c')$.

[11] Definition 2 looks similar to the Stationarity Postulate in Koopmans (1960, 1964) [9, 10]. However, time consistency and stationarity are conceptually very different (see below and Section 7.1).

[12] One could consider a weaker version of the second part of Definition 2: If $({}_{t+1}c) \succ^{t+1} ({}_{t+1}c')$, then $(c_t, {}_{t+1}c) \succsim^t (c_t, {}_{t+1}c')$. This version, however, would be at odds with time consistency, unless DM is fully myopic and cares only about the present.

[13] This step would be meaningless if $U({}_{t+1}c)$ represented how DM evaluates consumption streams starting at $t+1$ from the perspective of t, but not necessarily how he evaluates such streams from the perspective of $t+1$. This observation applies to the rest of the proof.

Suppose ${}_{t+1}c \sim {}_{t+1}c'$ implies $(c_t, {}_{t+1}c) \sim (c_t, {}_{t+1}c')$. Then, for any $(U({}_{t+1}c), U({}_{t+2}c), \ldots)$ and $(U({}_{t+1}c'), U({}_{t+2}c'), \ldots)$ such that $U({}_{t+1}c) = U({}_{t+1}c')$,

$$V(c_t, U({}_{t+1}c), U({}_{t+2}c), \ldots) = V(c_t, U({}_{t+1}c'), U({}_{t+2}c'), \ldots).$$

So can V depend only on its first two components. Suppose ${}_{t+1}c \succ {}_{t+1}c'$ implies $(c_t, {}_{t+1}c) \succ (c_t, {}_{t+1}c')$. Then, $U({}_{t+1}c) > U({}_{t+1}c')$. Moreover, it must be that $V(c_t, U({}_{t+1}c)) > V(c_t, U({}_{t+1}c'))$; that is, V must be strictly increasing in its second argument. □

As noted in the introduction, the standard DU model satisfies

$$U({}_tc) = u(c_t) + \delta U({}_{t+1}c) = V(c_t, U({}_{t+1}c)).$$

To allow for the two interpretations of this model mentioned in the introduction, this paper adopts the following definitions.

Definition 3 (Myopically Forward-Looking DM)**.** A DM is myopically forward-looking at t if either his immediate well-being depends only on his well-being at $t+1$, or he fails to realize that, in the future, he will continue to care about future selves.

Definition 4 (Fully Forward-Looking DM)**.** A DM is fully forward-looking if, at all $t \geq 0$, his immediate well-being depends directly on his well-being at all $s > t$.

To see the intuition behind Proposition 1, suppose ${}_{t+1}c$ involves reducing consumption (i.e., saving) at $t+1$—which DM dislikes—and increasing it at all following periods—which he likes. Also, suppose that, at $t+1$, DM is indifferent between saving immediately, ${}_{t+1}c$, or not, ${}_{t+1}c'$. Now imagine that, at t, DM can choose whether to save at $t+1$ or not, holding consumption at t fixed. Then, at t, if DM cares about his well-being beyond $t+1$, he must strictly prefer $(c_t, {}_{t+1}c)$ to $(c_t, {}_{t+1}c')$—i.e., saving at $t+1$. Intuitively, from the perspective of $t+1$, the negative effect on $t+1$ well-being of saving (i.e., of ${}_{t+1}c$) exactly offsets the positive effect of the higher well-being in later periods. But since at t DM cares *directly* about well-being beyond $t+1$, he weighs more the positive effect of ${}_{t+1}c$, thus strictly preferring to save at $t+1$. More generally, this mechanism can lead to situations in which, at $t+1$, DM wants to review the consumption plan chosen at t.

Proposition 1 has several implications. First, a DM who cares about his well-being beyond the immediate future must be time inconsistent. So this paper identifies a cause of time inconsistency in DM's caring, at t, about his well-being beyond $t+1$. Moreover, if we deem natural that a DM should care about his future in this way, then based on

Proposition 1, we must conclude that time inconsistency should be the rule, rather than the exception.

The second implication is that DM's preference is time consistent if and only if, at each t, it has a specific recursive representation: DM's utility from a consumption stream, $U({}_tc)$, depends only on immediate consumption, c_t, and continuation utility $U({}_{t+1}c)$. Moreover, at t, DM must always be happier if his continuation utility is higher. In this case, at $t+1$, DM never wants to review a consumption plan chosen at t, since that plan was already maximizing his $t+1$ well-being. In contexts in which DM at different dates correspond to different generations (e.g., as in Phelps-Pollack (1968) [18]), with this recursive utility each generation is benevolent towards its children, but cares about the well-being of its grandchildren and beyond only through the impact on its children's well-being. Put differently, this means that your grandparents care about you only because, if you are happy, then your parents are happy.

Definition 2 looks similar to the Stationarity Postulate in Koopmans (1960, 1964) [9, 10]. In the present setup, stationarity requires the following: If $({}_{t+1}c) \sim^t ({}_{t+1}c')$, then $(c_t, {}_{t+1}c) \sim^t (c_t, {}_{t+1}c')$; and if $({}_{t+1}c) \succ^t ({}_{t+1}c')$, then $(c_t, {}_{t+1}c) \succ^t (c_t, {}_{t+1}c')$. Despite the similarity, the two conditions are conceptually very different.[14] Indeed, Koopmans writes:

> "[Stationarity] does *not* imply that, after one period has elapsed, the ordering then applicable to the 'then' future will be the same as that now applicable to the 'present' future. All postulates are concerned with only one ordering, namely that guiding decisions to be taken in the present. Any question of change or consistency of preferences as the time of choice changes is therefore extraneous to the present study." (p. 85, [10], emphasis in the original)

The point can be illustrated with two simple examples.

Example 1 (Stationarity $\not\Rightarrow$ Time Consistency): For $t \geq 0$, $\succ^t$ is represented by

$$U_t({}_tc) = u(c_t) + \sum_{s \geq t+1}^{\infty} \delta_t^{s-t} u(c_s) = u(c_t) + \delta_t U_t({}_{t+1}c),$$

where $u(\cdot)$ is continuous, strictly increasing, and bounded. Moreover, $\delta_t \in (0,1)$ and $\delta_t > \delta_{t+1}$ for $t \geq 0$. Then, each $\succ^t$ satisfies stationarity. However, it is easy to see that $\{\succ^t\}_{t \geq 0}$ violates time consistency.

[14] Of course, if Assumption 1 holds, time consistency and stationarity are mathematically equivalent.

Example 2 (Time Consistency $\not\Rightarrow$ Stationarity): For $t \geq 0$, $\succ^t$ is represented by

$$U_t({}_tc) = u(c_t) + \phi_t \sum_{s \geq t+1}^{\infty} \delta^{s-t} u(c_s) = u(c_t) + \phi_t \delta U_{t+1}({}_{t+1}c),$$

where $\delta \in (0,1)$, $\phi_0 \in (0,1)$, and $\phi_t = 1$ for $t \geq 1$; $u(\cdot)$ has the same properties as before. It is easy to see that $\{\succ^t\}_{t\geq 0}$ satisfies time consistency, but $\succ^0$ violates stationarity.

These observations highlight in which way the present paper departs from other intertemporal utility models. It does not relax stationarity directly—for example, as Hayashi (2003) [8] and Olea and Strzalecki (2013) [16] do. Instead, it takes a conceptually different view of what determines preferences over consumption streams. The starting point is to allow forward-looking behavior to extend beyond the immediate future. By Proposition 1, doing so requires abandoning time consistency, which by time invariance (Assumption 1) happens to be formally equivalent to stationarity.

5 Discounting the Future: Well-being vs. Instantaneous Utility

Recursive utility representations are the norm in dynamic economic models, like repeated games or contracting games. By Proposition 1, however, such representations do not allow for fully forward-looking behavior. Standard geometric discounting of future instantaneous utility (not future well-being) is an example in this sense. More generally, let $d : \mathbb{N} \times \mathbb{N} \to \mathbb{R}_{++}$ be a discount function: $d(t,s)$ measures the weight that, sitting at t, DM assigns to date $s \geq t$. Suppose DM's preference are represented by

$$U({}_tc) = \sum_{s \geq t} d(t,s) u(c_s) \tag{3}$$

for $t \geq 0$, where $u : X \to \mathbb{R}$ is DM's instantaneous utility from consumption.[15] Without loss of generality, set $d(0,0) = 1$.

Proposition 2. *For $t \geq 0$, $U({}_tc) = u(c_t) + d(t,t+1)U({}_{t+1}c)$ if and only if $d(t,t) = 1$ and, for $s > t$,*

$$d(s,t) = \prod_{j=0}^{s-t-1} d(t+j, t+j+1). \tag{4}$$

[15] For simplicity, this paper focuses on the case of time-homogenous instantaneous utility.

The expression $U({}_tc) = u(c_t)+d(t,t+1)U({}_{t+1}c)$ is well known: It is the standard promise-keeping condition, where total utility today is the sum of instantaneous utility today plus the discounted continuation utility in the next period. It is also similar to a Bellman equation, where c_t is the optimal consumption at t.

Proposition 2 has two implications. First, if the discount function depends only on the time difference $s-t$, then we must have geometric discounting: $\hat{d}(s-t) = \delta^{s-t}$ for $\delta = \hat{d}(1)$. By Assumption 1 and Proposition 1, this of course says that geometric discounting of instantaneous utility implies time consistency. But this also says that any form of discounting that depends only on $s-t$ and is not geometric—like hyperbolic discounting—implies that, when comparing consumption streams, DM must care about his well-being beyond the immediate future. In other words, we can understand any departure from geometric discounting as a situation in which DM exhibits forward-looking behavior, as intended here.

The second implication concerns how future consumption affects current well-being. For $s > t$, this effect is captured by $d(t,s)$, which, by (4), takes account of only the direct channel through which consumption at s affects utility at t. As we will see, when DM cares about his well-being beyond next period, $d(t,s)$ will capture more than just this direct channel.

In light of Proposition 2, one wonders if it is possible to find a relationship between how DM discounts future instantaneous utility from consumption and how he cares about future well-being—for instance, when he discounts instantaneous utility hyperbolically, or quasi-hyperbolically. The next result relates instantaneous-utility discounting and well-being discounting. This result will help us understand the properties of how a forward-looking DM discounts consumption at each future time, from a more primitive well-being representation with discounting—which will be axiomatized in Section 7. To state the result, for $s > t$, let

$$\mathcal{T}(t,s) = \{\mathbf{t} = (\tau_0,\ldots,\tau_n) \mid 1 \le n \le |s-t|, \tau_0 = t, \tau_n = s, \tau_i < \tau_{i+1}\}, \tag{5}$$

and for $s \ge t+1$, let

$$\hat{\mathcal{T}}(t,s) = \{\mathbf{t} = (\tau_0,\ldots,\tau_n) \mid 2 \le n \le |s-t|, \tau_0 = t, \tau_n = s, \tau_i < \tau_{i+1}\}. \tag{6}$$

$\mathcal{T}(t,s)$ is the set of all increasing vectors starting at t and ending at s; $\hat{\mathcal{T}}(t,s)$ is the set of all such vectors with at least one intermediate date.

Proposition 3. *If* $U({}_tc) = u(c_t) + \sum_{s>t} q(t,s)U({}_sc)$ *for some well-being discount func-*

tion q, then $U({}_tc) = u(c_t) + \sum_{s>t} d(t,s)u(c_s)$ for some instantaneous-utility discount function d. Moreover, for all $t \geq 0$ and $s > t$,

$$d(t,s) = \sum_{\mathbf{t} \in \mathcal{T}(t,s)} \prod_{n=1}^{|\mathbf{t}|-1} q(\tau_{n-1}, \tau_n). \tag{7}$$

Similarly, if $U({}_tc) = u(c_t) + \sum_{s>t} d(t,s)u(c_s)$ for some instantaneous-utility discount function d, then $U({}_tc) = u(c_t) + \sum_{s>t} q(t,s)U({}_sc)$ for some (possibly negative) function q. In particular, for all $t \geq 0$ and $s > t+1$, $q(t,t+1) = d(t,t+1)$ and

$$q(t,s) = d(t,s) + \sum_{\mathbf{t} \in \hat{\mathcal{T}}(t,s)} (-1)^{|\mathbf{t}|} \prod_{n=1}^{|\mathbf{t}|-1} d(\tau_{n-1}, \tau_n). \tag{8}$$

The intuition behind Proposition 3 is simple. When DM is fully forward-looking, $d(t,s)$ takes account of all channels—not just the direct one as in Proposition 2—through which utility from consumption at s affects well-being at t (see (7)). There are as many such channels as there are ways of reaching s from t in j jumps, for $j \leq s-t-1$. For example, consumption two periods from now affects today's well-being via two channels: (1) a direct one, for today DM cares about his well-being two periods from now; (2) an indirect one, for today DM cares about his well-being tomorrow, which in turn depends on his well-being two periods from now. As a result, independently of how DM discounts well-being, he discounts instantaneous utility at least two periods in the future strictly less than his well-being at that period.

Corollary 1. *Suppose d and q are instantaneous-utility and well-being discount functions corresponding to the same preference $\succ$. Then, $d(t,s) > q(t,s)$ whenever $s - t \geq 2$.*[16]

As an illustration of the relationship between d and q, consider the well-known and empirically supported case of hyperbolic discounting (see Frederick et al. (2002) [6] and references therein). In its most general version, hyperbolic discounting takes the form

$$d(t,s) = d^h(s-t) = [1 + k(s-t)]^{-\frac{p}{k}} \tag{9}$$

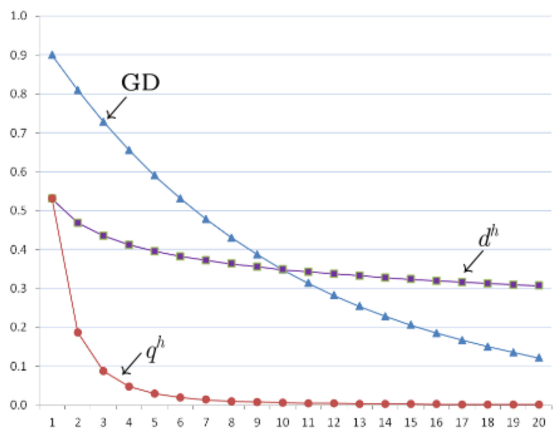
with k, $p > 0$ (see Loewenstein and Prelec (1992) [15]). Unfortunately, it is hard to derive in closed form the well-being discount function q^h corresponding to d^h.[17] However, using

[16]Using a recursive formulation, Saez-Marti and Weibull (2005) [20] make a similar but weaker observation, namely that $d(t,s) \geq q(t,s)$ for all t and s.

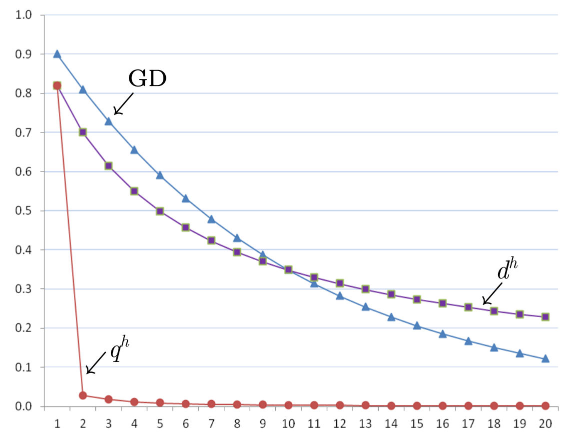
[17]Proposition 2 in Saez-Marti and Weibull (2005) [20] implies that, if d^h satisfies (9), then $q^h \geq 0$; that is, q^h is a discount function.

Proposition 3, we can simulate q^h for different values of k and p. This gives us an idea of how q^h qualitatively relates to d^h.

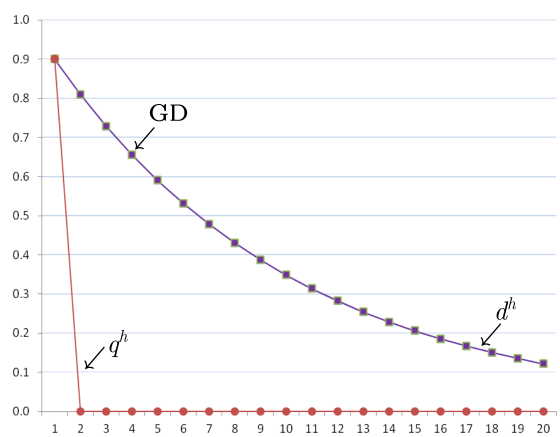
Figure 1: Hyperbolic-Discounting Simulation



(a) $k = 30$



(b) $k = 0.3$



(c) $k = 0.0001$

Figure 1 represents d^h, q^h, and the standard geometric discount function for $\delta = 0.9$ (curve GD). In all three panels, given k, the parameter p is set so that $d^h(\tau) = \delta^\tau$ at $\tau = 10$ (this period is arbitrary). Recall that hyperbolic discounting is characterized by declining discount rates, as highlighted by panel (a) and (b). Also, recall that, given p, $d^h(\tau)$ converges to δ^τ for every τ as $k \to 0$—this pattern is clear going from panel (a) to (c). All three panels illustrate the point of Corollary 1: Starting in period 2, q^h is strictly below d^h. Moreover, they illustrate that, when a hyperbolic-discounting time-inconsistent DM becomes more similar to a geometric-discounting time-consistent DM, he discounts more his well-being from period 2 onwards. Indeed, for these periods, q^h moves closer to zero as d^h moves closer to GD. This is, of course, an illustration of

Propositions 1 and 2: DM can be time consistent if and only if he does not care about well-being beyond period 1, i.e., $q^h(\tau) = 0$ for $\tau > 1$.

In general, taking account of all channels through which future consumption affects current well-being seems complicated. However, it can reduce to simple nongeometric discounting of instantaneous utility. The following result illustrates this point in the specific but important case of β-δ discounting.

Proposition 4. *Suppose that* $U({}_tc) = u(c_t) + \sum_{s>t} q(t,s)U({}_sc)$ *and* $U({}_tc) = u(c_t) + \sum_{s>t} d(t,s)u(c_s)$*. Then, for* $s > t$*,* $d(t,s) = \beta\delta^{s-t}$ *with* $\beta, \delta \in (0,1)$ *if and only if* $q(t,s) = \gamma\alpha^{s-t}$ *with* $\gamma > 0$ *and* $\alpha \in (0,1)$*. Moreover,*

$$\beta = \frac{\gamma}{1+\gamma} \text{ and } \delta = (1+\gamma)\alpha.$$

Section 7 presents a set of axioms on DM's preference that corresponds to a well-being representation of the form

$$V(c_t, U({}_{t+1}c), \ldots) = u(c_t) + \sum_{s>t} \alpha^{s-t}\gamma U({}_sc), \tag{10}$$

with $\gamma > 0$ and $\alpha \in (0,1)$ (see Theorem 4). In so doing, this paper offers a foundation of quasi-hyperbolic discounting based on the forward-looking behavior of a DM who takes account of his well-being beyond the immediate future. This interpretation differs from the usual one based on myopia, which says that a DM would discount future consumption quasi-hyperbolically because he disproportionately cares about the present against any future period. Moreover, Proposition 4 tightly links the 'degree of present bias,' β, with how DM 'rescales' with γ his future well-being before incorporating it in his current well-being. For example, geometric discounting of well-being ($\gamma = 1$) corresponds to $\beta = \frac{1}{2}$. We can interpret the rescaling parameter γ as measuring the degree to which DM finds future outcomes 'imaginable' or 'vivid.' As suggested by Böhm-Bawerk (1889) [24] and Fisher (1930) [5], DM's current utility depends on current consumption as well as on his ability to imagine or foresee his future 'wants.' This essential difference between current consumption and future well-being results in an asymmetric treatment of current and future instantaneous utilities from consumption, which are instead identical in nature. Finally, note that the representation in (10) separates 'vividness' of future well-being from discounting.

Proposition 4 has several implications. First, it implies that fully forward-looking behavior inevitably leads to present bias with respect to instantaneous utility—indeed, the implied β is always strictly less than one. Moreover, it can also give an explanation for why present bias weakens as the period length shortens. If each period t represents a shorter time horizon, current instantaneous utility $u(c_t)$ should play a smaller role in determining current well-being; that is, γ should be larger in (10). Consequently, β should get closer to one. The second implication is about long-run rates of discounting and how increasing the vividness of future well-being can mitigate present bias.

Corollary 2. *For any degree of present bias β, a fully forward-looking DM discounts instantaneous utility at a long-run rate δ that is strictly higher than the rate α at which he discounts well-being. Moreover, for any α, increasing vividness, γ, mitigates present bias and improves the long-run rate δ.*

For example, using Laibson et al.'s (2007) [13] estimates of $\beta = 0.7$ and $\delta = 0.95$, we get $\gamma = 2.33$ and $\alpha = 0.285$. So well-being discounting is actually much steeper than one might think by looking only at consumption discounting; moreover, this difference depends significantly on 'vividness.'

If we interpret DM in different periods as different generations, then Corollary 2 suggests that fully forward-looking generations prefer policies that are 'sustainable' in the long run. If c and c' involve long-run decisions that yield the same well-being to the next generation, but c' harms all later generations more than c, then the current generation would prefer c to c'; in contrast, with a recursive utility function, the current generation would be indifferent between c and c'. Of course, fully forward-looking generations are also present biased, so they are subject to procrastinating the implementation of c'—e.g., curbing carbon emissions. By the second part of Corollary 2, however, increasing the vividness of the consequences of c and c' on future generations both strengthens the preference for sustainability and mitigates procrastination.[18]

Finally, one can use Proposition 4 to see that a fully forward-looking DM can exhibit a desire to advance dreadful events and postpone delightful ones, or a preference for improving (worsening) sequences of good (bad) events. Evidence of these phenomena, which are at odds with standard discounted utility, is summarized in Frederick et al. (2002) [6]. In Proposition 4, δ must be strictly less than one to ensure that U is well defined over infinite-horizon problems. However, with finite horizon, δ can exceed 1, even if DM positively discounts well-being ($\alpha < 1$); that is, DM can weight future instantaneous utility *more* than the current one, even if he weights future well-being *less* than the current one.

[18] This assumes that external forces—like government ad campaigns—can modify DM's preference, in the spirit of the literature that interprets advertising as a way of modifying people's preferences.

Example 1. (Anticipating Dreadful Events) DM has to undergo a colonoscopy, denoted by x. He finds this treatment very vivid (γ large) and painful ($u(x) < 0$). Recall that $\beta\delta = \gamma\alpha$. So, if $\delta > 1$, DM strictly prefers to have his colonoscopy either today (if $\beta\delta > 1$), or tomorrow (if $\beta\delta < 1$), than at any future date.

Example 2. (Postponing Delightful Events) DM has to choose when to open a very expensive bottle of wine with his friends, c_t. Based on past experience, DM finds such events vivid (γ large) and enjoyable ($u(c_t) > 0$). Then, if $\beta\delta < 1$, DM would prefer to open the bottle today rather than tomorrow. But if $\delta > 1$, DM may strictly prefer to open it at some time after tomorrow—provided that the natural decay of the wine quality is not too fast nor too slow.[19]

More generally, it is easy to see that $\delta > 1$ can lead DM to prefer improving (rather than worsening) sequences of consumption, in a way that contradicts the predictions of standard discounted utility but is consistent with some empirical evidence (see, e.g., Frederick (2002) [6]). If $\delta > 1$, at t, DM assigns increasing weights to instantaneous utilities from $t + 1$ onward. Therefore, given consumption at t, he is better off choosing sequences that assign increasing, rather than decreasing, levels of consumption (and instantaneous utility) to future periods.

6 Discussion: Welfare Criteria and Normative Analysis

Models that allow for time-inconsistent preferences pose serious conceptual problems, when defining welfare criteria and addressing policy questions. Discussing hyperbolic discounting, Rubinstein (2003) [19] notes:

> "Policy questions were freely discussed in these models even though welfare assessment is particularly tricky in the presence of time inconsistency. The literature often assumed, though with some hesitation, that the welfare criterion is the utility function with stationary discounting rate δ (which is independent of β)." (p. 1208)

Alternatively, to assess welfare, one may want to use a different weighted sum of each-period instantaneous utilities, possibly averaging the β-δ weights with the standard DU weights.

[19]For instance, suppose the wine goes bad at some $t' < +\infty$, so that $u(c_t) = 0$ for $t \geq t'$.

The present paper offers a different, perhaps clearer, interpretation of welfare with time-inconsistent preferences. Since well-being representations depend explicitly on DM's well-being at each date, they allows us to understand how current well-being takes into account future well-being. In the specific case of the β-δ model, for instance, the parameters β and δ should be mapped back to the parameters γ and α in the well-being representation, as indicated by Proposition 4. With this representation, a level of well-being can be explicitly assigned to each DM's self (or generation).

This last point highlights that how a welfare criterion should weigh each-period well-being is a more general problem and has nothing to do with (quasi-) hyperbolic discounting or, more generally, with time inconsistency. Indeed, in the standard DU model, DM's welfare is usually and uncontroversially measured as the sum of instantaneous utilities, discounted at DM's subjective rate. That model, however, implies that DM's current preference takes into account only his well-being in the next period, but not in future periods. In contrast, the β-δ model implies that DM's current preference takes into account, though in a simple way, his well-being in all future periods—the same holds with more general well-being representations. From this perspective, the ex-ante preference of the β-δ model appear as a more reasonable welfare criterion, than the ex-ante preference of a DU model with discount rate δ.

To see this point from a different angle, consider a model with an infinite sequence of generations, in which each generation's well-being is measured by the geometrically discounted sum of future generations' instantaneous utility. It is standard to use such a well-being of the first generation, so as to evaluate the welfare implied by different policies. The first generation, however, does not care directly about the well-being of its grandchildren, great grandchildren, and so forth.

This property seems to call for a different welfare criterion. In the present model, each generation cares directly about the well-being of future generations beyond its children, and its choices reveal how much. For example, we could interpret α^t in Proposition 4 as representing how much generation 0 cares about generation t, whose well-being has vividness γ for generation 0. A planner may then use this information to build a welfare criterion for all generations. On the one hand, the planner may simply weigh unborn generations in the same way as does the first generation. On the other hand, the planner may, for example, aggregate the well-being of all generations using weights α^t, so as to determine how to weigh each generation's consumption.

7 Axiomatization

7.1 General Well-Being Representation

This section first presents a set of axioms that corresponds to DM's preference having the general well-being representation in Definition 1. With further axioms, it then obtains the more specific forms analyzed in Section 4. Since DM's preference are time invariant (Assumption 1), for simplicity, this section takes the perspective of $t = 0$. Hereafter, the set X of feasible consumption levels is a connected, separable metric space with metric d. Let $>$ be an order on X—e.g., if $X \subset \mathbb{R}$, $>$ is the usual order. The set $C = X^T$ of consumption streams is endowed with the sup-norm: $\|c - c'\|_C = \sup_t d(c_t, c'_t)$.

The first three axioms are standard.

Axiom 1 (Weak Order). *$\succ$ is complete and transitive.*

Axiom 2 (Continuity). *For all $c \in \mathcal{C}$, the sets $\{c' \in \mathcal{C} : c' \prec c\}$ and $\{c' \in \mathcal{C} : c' \succ c\}$ are open.*

Axiom 3 (Constant-Flow Dominance). *For all $c \in \mathcal{C}$, there exist constant sequences $\overline{c}$ and $\underline{c}$ such that $\underline{c} \precsim c \precsim \overline{c}$.*

These axioms lead to the following standard result, which builds on Diamond (1965) [4].

Theorem 1. *Under Axioms 1-3, there exists a continuous function $U : \mathcal{C} \to \mathbb{R}$ that represents $\succ$.*

Theorem 1 says that DM's preference at each t can be expressed in terms of current and future consumption, as in a standard model. The function U obtained here, however, will be different from the one arising under usual assumptions: $U(c)$ assumes the connotation of the well-being generated by consumption stream c, which depends on immediate consumption as well as future well-being.

The following axiom simply rules out a trivial dependence on immediate consumption as well as full myopia.

Axiom 4 (Non triviality). *There exist $x, x', \hat{x} \in X$ and ${}_1c, {}_1c', {}_1\hat{c} \in {}_1\mathcal{C}$, such that*

$$(x, {}_1\hat{c}) \succ (x', {}_1\hat{c}) \text{ and } (\hat{x}, {}_1c) \succ (\hat{x}, {}_1c').$$

The next axiom aims to capture the idea that consumption streams starting tomorrow affect well-being today, only through the well-being that they generate at each future time. This axiom is the key to obtaining the well-being representation in Definition 1.

Axiom 5. *If ${}_tc \sim {}_tc'$ for all $t \geq 1$, then $(c_0, {}_1c) \sim (c_0, {}_1c')$.*

Axiom 5 rules out the possibility that DM prefers stream c over c' because, even though they generate the same stream of immediate consumption and future well-being, they allocate future consumption differently over time. To express DM's preference in terms of immediate consumption and future well-being, for $t \geq 1$, let $f_0(c) = c_0$, $f_t(c) = U({}_tc)$, and $f(c) = (f_0, f_1, f_2, \ldots)$. Also, let

$$\mathcal{F} = \{f(c) : c \in \mathcal{C}\}, \tag{11}$$

and let $\mathcal{U}$ be the range of U. Note that $\mathcal{U}$ is an interval by continuity of U, non triviality (Axiom 4), and connectedness of X.

Theorem 2. *Axioms 1-5 hold if and only if there exists a continuous function $V : X \times \mathcal{U}^{\mathbb{N}} \to \mathbb{R}$ such that $V(f(c)) = U(c)$ and V is nonconstant in f_0 and f_t for some $t \geq 1$.*

Proof. ($\Rightarrow$) First, define equivalence classes on consumption streams as follows. Say that c is equivalent to c' if $U({}_tc) = U({}_tc')$ for all $t \geq 1$ and $c_0 = c'_0$.[20] Let $\mathcal{C}^*$ be the set of equivalence classes of $\mathcal{C}$, and let U^* be defined by the utility function U on that domain. Then, the function $f^* : \mathcal{C}^* \to \mathcal{F}$, defined by $f^*(c^*) = f(c)$ for c in the equivalence class c^*, is by construction one-to-one and onto; so let $(f^*)^{-1}$ denote its inverse. Finally, for any $f \in \mathcal{F}$, define

$$V(f) = U^*((f^*)^{-1}(f)).$$

By Axiom 5, V is a well-defined function, and $V(f(c)) = U(c)$ for every c. By Axiom 4, V is nonconstant in its first argument; moreover, since $V(f(\hat{x},_1 c)) > V(f(\hat{x},_1 c'))$, the function V cannot be constant in all arguments other than the first one.

($\Leftarrow$) Suppose $V : \mathcal{F} \to \mathbb{R}$ is a continuous function such that $V(f(c)) = U(c)$. Then, it is immediate to see that the implied preference $\succ$ satisfies Axioms 1-4. $\square$

Proposition 2 allows us to represent DM's preference in terms of current consumption, f_0, and future well-being, f_t for $t \geq 1$. Note that Axiom 5 is weaker than an axiom—

[20] In general, there may be several consumption streams in an equivalence class. For example, suppose $U(c) = c_0 + c_1 + c_2 + c_3$, and let $c = (1, 1, -1, -1, 1, 1, -1, -1, \ldots)$ and $c' = (1, -1, -1, 1, 1, -1, -1, 1, 1, \ldots)$.

similar, in spirit, to Koopmans (1960) [9] stationarity—that requires ${}_1c \sim {}_1c'$ to imply $(c_0, {}_1c) \sim (c_0, {}_1c')$. Axiom 5 requires that DM be indifferent (today) between two consumption streams, only if he is indifferent between their truncations at *all* future times, not just tomorrow.

7.2 Discounting Well-Being Representations

This section aims to refine the general well-being representation V (Definition 1), by introducing time separability and a stationary dependence on future well-being.

The first two axioms, inspired by Debreu (1960) [3] and Koopmans (1960) [9], imply that DM's preference are separable in immediate consumption and future well-being, as well as across future well-being. Let Π consist of all unions of subsets of $\{\{1\}, \{2\}, \{3, 4, \ldots\}\}$.

Axiom 6 (Immediate-Consumption and Well-Being Separability). *Fix any* $\pi \in \Pi$. *If* c, $\hat{c}$, c', $\hat{c}' \in \mathcal{C}$ *satisfy*
- ${}_tc \sim {}_t\hat{c}$ *and* ${}_tc' \sim {}_t\hat{c}'$ *for all* $t \in \pi$,
- ${}_tc \sim {}_tc'$ *and* ${}_t\hat{c} \sim {}_t\hat{c}'$ *for all* $t \in \mathbb{N} \setminus \pi$,
- *either* $c_0 = c'_0$ *and* $\hat{c}_0 = \hat{c}'_0$, *or* $c_0 = \hat{c}_0$ *and* $c'_0 = \hat{c}'_0$,

then $c \succ c'$ *if and only if* $\hat{c} \succ \hat{c}'$.

Note that Axiom 6 requires that certain consumption streams be indifferent, as opposed to being equal. This is because we want separability in well-being, which can be the same across streams that allocate consumption differently over time. The next axiom is of technical nature. It ensures that DM's preference does depend on well-being at period 1, 2, and 3 (Debreu's essentiality condition).

Axiom 7 (Essentiality). *There exist* x, x', y, $y' \in X$ *and* ${}_3c$, ${}_3c' \in {}_3\mathcal{C}$ *such that* $(z, x,_2 c) \succ (z, x',_2 c)$, $(z', z'', y,_3 c'') \succ (z', z'', y',_3 c'')$, *and* $(w, w', w'',_3 c) \succ (w, w', w'',_3 c')$ *for some* z, z', z'', w, w', $w'' \in X$, ${}_2c$ *and* ${}_3c''$.

The third axiom, also inspired by Koopmans (1960) [9], is meant to ensure that DM's preference is sufficiently stationary. Intuitively, stationarity means that, from today's point of view ($t = 0$), DM does not see himself as changing how he evaluates an event, simply because this event is postponed to a subsequent date. Of course, requiring stationarity is reasonable only to the extent that this property refers to postponing events or objects of the same nature. In this paper, however, instantaneous consumption and future well-being are conceptually different. So the axiom requires stationarity only with respect to future well-being.

Axiom 8 (Well-Being Stationarity). *If c, $c' \in \mathcal{C}$ satisfy $c_0 = c'_0$ and ${}_1c \sim {}_1c'$, then*

$$(c_0, {}_2c) \succsim (c'_0, {}_2c') \iff c \succsim c'.$$

It is useful to compare this axiom with Koopmans' (1960) [9] stationarity and with Olea and Strzalecki's (2013) quasistationarity. Koopmans' stationarity involves postponing two consumption streams from today to tomorrow, replacing today's consumption with the same amount in both cases. Stationarity requires that DM rank such new streams as he ranked the original streams. If we assume that DM evaluates streams based only on the instantaneous utility generated in each period—so that today's and tomorrow's consumption are conceptually equivalent—then stationarity seems reasonable (at least normatively).

On the other hand, Olea and Strzalecki allow for violations of stationarity, but assume quasistationarity. This property involves postponing two consumption streams (with equal today's consumption) from tomorrow to the day after, replacing tomorrow's consumption with the same *amount* in both cases. Again, quasistationarity requires that DM rank such new streams as he ranked the original ones. Olea and Strzalecki continue to assume that DM evaluates streams based only on instantaneous utilities. However, within this framework, quasistationarity seems more difficult to justify. If DM views consumption in the same way in all periods, why should stationarity hold between tomorrow and the day after, but not between today and tomorrow? This issue does not arise in the present model, for tomorrow's well-being is equivalent to well-being thereafter, but differs conceptually from today's consumption.

Finally, the next axiom says that DM likes consumption and is benevolent towards his future selves (or generations).

Axiom 9 (Monotonicity). *If c, $c' \in \mathcal{C}$ satisfy $c_0 > c'_0$ and ${}_tc \succsim {}_tc'$ for all $t \geq 1$, then $c \succsim c'$.*

Theorem 3 (Additive Well-Being Representation). *Axiom 1-8 hold if and only if there exist continuous nonconstant functions $u : X \to \mathbb{R}$ and $G : \mathcal{U} \to \mathbb{R}$, and $\alpha \in (0, 1)$ such that*

$$U(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U({}_tc)).$$

Moreover, Axiom 9 also holds if and only if u and G are also increasing.

One might wonder whether the cumulative effects of future well-being on preceding well-being inevitably leads the summation in the expression of Theorem 3 to diverge to infinity. By Axioms 1-4, the function U is a well-defined representation of $\succ$ and cannot be constant. Therefore, there always exist consumption streams c such that $U(c)$ is bounded. Of course, if U is always bounded, then the summation in the expression of Theorem 3 will always converge.

By Axiom 6, separability holds between immediate consumption and future well-being, as well as across future well-being. Nonetheless, since well-being at t affects well-being at all $s < t$, the trade-off between consumption at 0 and at t can depend on well-being—and hence consumption—between 0 and t. So, controlling for the effect of instantaneous utility at 0 and t, we get that DM's discount factor can depend on well-being between 0 and t. Of course, this dependence hinges on the properties of G. On the other hand, well-being at t—hence consumption at t and thereafter—can affect the trade-off between consumption at any two preceding dates.[21] For example, note that

$$\begin{aligned} u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U({}_tc)) &= u(c_0) + \alpha G\left(u(c_1) + \sum_{t=2}^{\infty} \alpha^t G(U({}_tc)) \right) \\ &\quad + \sum_{t=2}^{\infty} \alpha^t G(U({}_tc)); \end{aligned}$$

so well-being at all dates after $t = 2$ can affect the trade-off between consumption today and tomorrow. This effect, however, is absent in the standard discounted-utility model, which highlights that Axiom 6 is weaker than the usual separability axiom.

More generally, the previous effect would be absent if G were affine, a property that is implied by the next axiom. As a result, we obtain a representation in terms of 'vividness' of future well-being and, in light of Proposition 4, an axiomatic foundation for the renowned β-δ discounting of instantaneous utility.

Axiom 10 (Strong Consumption Separability)**.**
i) $(c_0, c_1, {}_2c) \succ (c_0, c_1', {}_2c) \iff (\hat{c}_0, c_1, {}_2\hat{c}) \succ (\hat{c}_0, c_1', {}_2\hat{c})$
ii) $(c_0, c_1, {}_2c) \succ (c_0', c_1, {}_2c') \iff (c_0, \hat{c}_1, {}_2c) \succ (c_0', \hat{c}_1, {}_2c')$
iii) $(c_0, c_1, {}_2c) \succ (c_0', c_1', {}_2c) \iff (c_0, c_1, {}_2\hat{c}) \succ (c_0', c_1', {}_2\hat{c})$
iv) $(c_0, c_1, {}_2c) \succ (c_0, c_1, {}_2c') \iff (\hat{c}_0, \hat{c}_1, {}_2c) \succ (\hat{c}_0, \hat{c}_1, {}_2c')$.

Axiom 10 and axiom 6 with $\pi = \varnothing$ imply strong separability in c_0, c_1, and ${}_2c$. Note, however, that these two axioms are quite different: Axiom 6 is about separability in well-

[21] For evidence of this phenomenon see, e.g., Frederick et al. (2002) [6].

being, whereas axiom 10 is about separability in consumption in the first two periods and well-being in the third period.

Theorem 4 ('Vividness' Well-being Representation)**.** *Axiom 1-10 hold if and only if there exist $\gamma > 0$, $\alpha \in (0,1)$, and a continuous nonconstant function u such that*

$$U(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t \gamma U({}_tc).$$

This result allows us to understand β-δ discounting of instantaneous utility in terms of properties of DM's preference for immediate consumption and future well-being. First, the stark, apparently ad-hoc, difference between discounting of future instantaneous utility from today's perspective and from any future period's perspective comes from the natural conceptual difference between today's utility from immediate consumption and future well-being. Second, additive time separability in instantaneous utility, which characterizes β-δ discounting, requires separability in immediate consumption and future well-being, but also that the trade-off between consumptions at any two consecutive periods is not affected by the following well-being.

8 Appendix

8.1 Proof of Proposition 2

($\Rightarrow$) Rewrite (3) as

$$\begin{aligned} U({}_tc) &= d(t,t)u(c_t) + \sum_{s>t} d(t,s)u(c_s) \\ &= d(t,t)u(c_t) + d(t,t+1)\left[u(c_{t+1}) + \sum_{s>t+1} \frac{d(t,s)}{d(t,t+1)} u(c_s)\right], \end{aligned} \tag{12}$$

and

$$U({}_{t+1}c) = d(t+1,t+1)u(c_{t+1}) + \sum_{s>t+1} d(t+1,s)u(c_s). \tag{13}$$

For $U({}_tc) = u(c_t) + d(t,t+1)U({}_{t+1}c)$ to hold for all $t \geq 0$, it must be that $d(t,t) = 1$ for all $t \geq 1$, and

$$d(t,s) = d(t,t+1)d(t+1,s) \tag{14}$$

for all $t \geq 0$ and $s > t+1$.[22] In particular, it follows that $d(t,t+2) = d(t,t+1)d(t+1,t+2)$

[22]For this procedure to be justified, one needs enough variability in $u(c)$ across consumption, which is

for all $t \geq 0$. Suppose that, for all $t \geq 0$ and $2 \leq k \leq n-1$,

$$d(t, t+k) = \prod_{j=0}^{k-1} d(t+j, t+j+1).$$

Then, by (14),

$$\begin{aligned} d(t, t+n) &= d(t, t+1)d(t+1, t+n) \\ &= d(t, t+1)d(t+1, (t+1)+n-1) \\ &= d(t, t+1) \prod_{j=0}^{(n-1)-1} d((t+1)+j, (t+1)+j+1) \\ &= \prod_{j=0}^{n-1} d(t+j, t+j+1). \end{aligned}$$

The result follows by induction.

($\Leftarrow$) It follows by substituting $d(t, s)$ into (12) and (13).

8.2 Proof of Proposition 3

(Part 1) Suppose $\succ$ can be represented by $U({}_tc) = u(c_t) + \sum_{s>t} q(t, s)U({}_sc)$. We want to show that there exists an alternative representation of $\succ$, given by

$$U({}_tc) = \sum_{s \geq t} d(t, s)u(c_s)$$

for some discount function d. If this is true, then

$$\begin{aligned} U({}_tc) &= u(c_t) + \sum_{s>t} q(t, s)U({}_sc) \\ &= u(c_t) + \sum_{s>t} q(t, s) \left[\sum_{r \geq s} d(s, r)u(c_r) \right] \\ &= u(c_t) + \sum_{s>t} \left[\sum_{t<r \leq s} q(t, r)d(r, s) \right] u(c_s). \end{aligned}$$

assumed here. For example, if u is identically equal to zero, then no identification of the coefficients is feasible. Instead, if u takes at least two value over X, then the identification method is justified: Varying consumption at a single t—leaving the rest of consumption unchanged—pins down the coefficient at t.

This implies that, for all $t \geq 0$ and $s > t$, $d(t,t) = 1$ and

$$d(t,s) = \sum_{t<r\leq s} q(t,r)d(r,s). \tag{15}$$

This immediately implies $d(t,t+1) = q(t,t+1)$ for all $t \geq 0$. Now define $\mathcal{T}(t,s)$ as in (5). Suppose that (7) holds for all $t \geq 0$ and $s = t+k$ with $1 \leq k \leq n-1$. Then, by (15)

$$\begin{aligned} d(t,t+n) &= \sum_{t<r\leq t+n} q(t,r)d(r,s) \\ &= q(t,t+n) + \sum_{t<r\leq t+n-1}\left[\sum_{\mathbf{t}\in\mathcal{T}(r,t+n)} \prod_{j=1}^{|\mathbf{t}|-1} q(\tau_{j-1},\tau_j)q(t,r)\right] \\ &= q(t,t+n) + \sum_{\mathbf{t}\in\mathcal{T}(t,t+n)\setminus\{(t,t+n)\}} \prod_{j=1}^{|\mathbf{t}|-1} q(\tau_{j-1},\tau_j)q(t,r). \end{aligned}$$

The result follows by induction.

(Part 2) Suppose $\succ$ can be represented by $U({}_tc) = u(c_t) + \sum_{s>t} d(t,s)u(c_s)$. We want to show that there exists an alternative representation of $\succ$, given by

$$U({}_tc) = u(c_t) + \sum_{s>t} q(t,s)U({}_sc),$$

for some function q. If this is true, then for all $t \geq 0$,

$$u(c_t) = U({}_tc) - \sum_{s>t} q(t,s)U({}_sc)$$

and

$$\begin{aligned} U({}_tc) &= u(c_t) + \sum_{s>t} d(t,s)\left[U({}_sc) - \sum_{r>s} q(s,r)U({}_rc)\right] \\ &= u(c_t) + d(t,t+1)U({}_{t+1}c) + \sum_{s>t+1} U({}_sc)\left[d(t,s) - \sum_{t<r\leq s-1} q(r,s)d(t,r)\right] \end{aligned}$$

This implies that, for all $t \geq 0$ and $s > t+1$, $q(t,t+1) = d(t,t+1)$ and

$$q(t,s) = d(t,s) - \sum_{t<r\leq s-1} q(r,s)d(t,r). \tag{16}$$

Define $\hat{\mathcal{T}}(t,s)$ as in (6). For $s=t+2$, (16) becomes

$$\begin{aligned} q(t,t+2) &= d(t,t+2)-d(t,t+1)q(t+1,t+2) \\ &= d(t,t+2)-d(t,t+1)d(t+1,t+2). \end{aligned}$$

So (8) holds for all $t \geq 0$ and $s=t+2$, since $\hat{\mathcal{T}}(t,t+2)=\{(t,t+1,t+2)\}$. Now suppose that (8) holds for all $t \geq 0$ and $s=t+k$ with $2 \leq k \leq n-1$. Then, by (16), for $s'=t+n$

$$\begin{aligned} q(t,s') &= d(t,s')-d(t,s'-1)d(s'-1,s') \\ &\quad - \sum_{t<r\leq s'-2} d(t,r)\left[d(r,s')+\sum_{\mathbf{t}\in\hat{\mathcal{T}}(r,s')}(-1)^{|\mathbf{t}|}\prod_{j=1}^{|\mathbf{t}|-1} d(\tau_{j-1},\tau_j)\right] \\ &= d(t,s')-\sum_{t<r\leq s'-1} d(r,s')d(t,r)+ \\ &\quad -\sum_{t<r\leq s'-2}\sum_{\mathbf{t}\in\hat{\mathcal{T}}(r,s')}(-1)^{|\mathbf{t}|}\prod_{j=1}^{|\mathbf{t}|-1} d(\tau_{j-1},\tau_j)d(t,r) \\ &= d(t,s')+\sum_{\mathbf{t}\in\hat{\mathcal{T}}(t,s')}(-1)^{|\mathbf{t}|}\prod_{j=1}^{|\mathbf{t}|-1} d(\tau_{j-1},\tau_j). \end{aligned}$$

The result follows by induction.

8.3 Proof of Proposition 4

($\Leftarrow$) Suppose $q(t,s)=\gamma\alpha^{s-t}$ with $\gamma>0$ and $\alpha\in(0,1)$. By (7) of Proposition 3, for all $t\geq 0$ and $s>t$,

$$d(t,s)=\sum_{\mathbf{t}\in\mathcal{T}(t,s)}\prod_{j=1}^{|\mathbf{t}|-1}\gamma\alpha^{\tau_j-\tau_{j-1}}=\alpha^{s-t}\sum_{\mathbf{t}\in\mathcal{T}(t,s)}\gamma^{|\mathbf{t}|-1}.$$

The second factor on the right can be written as

$$\sum_{i=0}^{s-t-1}\tbinom{s-t-1}{i}\gamma^{i+1}. \tag{17}$$

Indeed, for each $0\leq i\leq s-t-1$, there are $\binom{s-t-1}{i}$ ways to go from t to s in $i+1$ jumps, because there are $\binom{s-t-1}{i}$ ways to choose the $i+1$ locations of the jumps. Expression (17) equals $\gamma(1+\gamma)^{s-t-1}$, which can be easily seen from the binomial development of $(1+x)^n$. This implies that

$$d(t,s)=\frac{\gamma}{1+\gamma}((1+\gamma)\alpha)^{s-t}.$$

So $\beta = \frac{\gamma}{1+\gamma}$, which is in $(0,1)$ since $\gamma > 0$, and $\delta = (1+\gamma)\alpha > 0$ since $\gamma, \alpha > 0$. To see why $\delta < 1$, consider $x \in X$ such that $|u(x)| \neq 0$. Then,

$$U(x, x, \ldots) = u(x) \left(1 + \beta \sum_{s>0}^{\infty} \delta^s\right),$$

so $\delta < 1$ because U is bounded.

($\Rightarrow$) Suppose $d(t,s) = \beta\delta^{s-t}$ with $\beta \in (0,1)$ and $\delta > 0$. By (8) of Proposition 3, for all $t \geq 0$ and $s > t+1$,

$$\begin{aligned} q(t,s) &= \beta\delta^{s-t} + \sum_{\mathbf{t} \in \hat{\mathcal{T}}(t,s)} (-1)^{|\mathbf{t}|} \prod_{n=1}^{|\mathbf{t}|-1} \beta\delta^{\tau_n - \tau_{n-1}} \\ &= \delta^{s-t} \left[\beta - \sum_{\mathbf{t} \in \hat{\mathcal{T}}(t,st)} (-\beta)^{|\mathbf{t}|-1}\right] \\ &= \delta^{s-t} \left[\beta - \sum_{j=1}^{s-t-1} \binom{s-t-1}{j} (-\beta)^{j+1}\right] \\ &= \delta^{s-t}\beta \sum_{j=0}^{s-t-1} \binom{s-t-1}{j} (-\beta)^j = \frac{\beta}{1-\beta}((1-\beta)\delta)^{s-t}. \end{aligned}$$

In the third line, the sum starts at $j = 1$ because, by definition, $\hat{\mathcal{T}}(t,s)$ contains vectors with at least two jumps. Finally, $q(t,t+1) = d(t,t+1) = \frac{\beta}{1-\beta}((1-\beta)\delta)$.

8.4 Proof of Theorem 1

The proof follows and generalizes that of Diamond (1965) [4], and is based on the following lemma by Debreu (1954) [2].

Lemma 1. *Let $\mathcal{C}$ be a completely ordered set and $Z = (z_0, z_1, ...)$ be a countable subset of $\mathcal{C}$. If for every pair c, c' of elements of $\mathcal{C}$ such that $c \prec c'$, there is an element z of Z such that $c \preceq z \preceq c'$, then there exists on $\mathcal{C}$ a real, order-preserving function, continuous in any natural topology.*[23]

Lemma 2. *For any $c \in \mathcal{C}$, there exists a constant sequence c^* such that $c \sim c^*$.*

Proof. Let $\mathcal{D}$ be the set of constant sequences and, for any fixed $c \in \mathcal{C}$, let $A = \{d \in \mathcal{D} : d \precsim c\}$ and $B = \{d \in \mathcal{D} : d \succsim c\}$. Axiom 2 implies that A and B are nonempty and closed; Axiom 1 implies that $A \cup B = \mathcal{D}$. Moreover, $\mathcal{D}$ is connected. Indeed, for any continuous

[23] A natural topology is one under which Axiom 2 holds for that topology.

function $\phi : \mathcal{D} \to \{0,1\}$, the function $\bar{\phi} : X \to \{0,1\}$ defined by $\bar{\phi}(x) = \phi(x, x, \ldots)$ is also continuous. Connectedness of X implies that $\bar{\phi}$ is constant and, hence, that ϕ is constant, showing connectedness of $\mathcal{D}$. This implies that A and B have a nonempty intersection, proving the lemma.

□

To conclude the proof of Theorem 1, let Z_0 denote a countable subset of X, which exists since X is separable, and let Z denote the subset $\mathcal{C}$ consisting of constant sequences whose elements belong to Z_0. Lemma 2 implies that Z satisfies the hypothesis of Lemma 1, which yields the result.

8.5 Proof of Theorem 3

This proof adapts arguments in Debreu (1960) [3] and Koopmans (1960, 1964) [9, 10] to the present environment. It is convenient to work in terms of the streams of immediate consumption and future well-being f, defined in (11), and the binary relation $\succ^*$ on $\mathcal{F}$ induced by the function $V : \mathcal{F} \to \mathbb{R}$ in the proof of Theorem 2.

Let Π' consist of all unions of subsets of $\{\{0\}, \{1\}, \{2\}, \{3, 4, \ldots\}\}$.

Lemma 3. *Axiom 6 implies that $\succ^*$ satisfies the following property. For any f, $f' \in \mathcal{F}$ and $\pi \in \Pi'$,*

$$(f_\pi, f_{\pi^c}) \succ^* (f'_\pi, f_{\pi^c}) \iff (f_\pi, f'_{\pi^c}) \succ^* (f'_\pi, f'_{\pi^c}),$$

where $\pi^c = T \setminus \pi$. By Axiom 7, $\succ^*$ depends on f_0, f_1, f_2, and ${}_3f$.

Proof. Recall that ${}_tc \sim {}_tc'$ implies $U({}_tc) = U({}_tc')$, which is equivalent to $f_t = f'_t$. Then, by Axiom 6, for any $\pi \in \Pi'$

$$V(f_\pi, f_{\pi^c}) > V(f'_\pi, f_{\pi^c}) \iff V(f_\pi, f'_{\pi^c}) > V(f'_\pi, f'_{\pi^c}).$$

□

By Debreu [3], there exist then continuous nonconstant functions $\overline{V}$, $\hat{u}$, a, b, and d such that

$$\overline{V}(f) = \hat{u}(f_0) + a(f_1) + b(f_2) + d({}_3f) \tag{18}$$

and

$$f \succ^* f' \iff \overline{V}(f) > \overline{V}(f').$$

By Lemma 3 with $\pi = \{0\}$, Axiom 4, and Koopmans' [9] argument, V can be expressed as

$$V(f) = W(v(f_0), A({}_1f)) \tag{19}$$

for some continuous, nonconstant functions W, v, and A. Similarly, by Lemma 3 with $\pi = \{1\}$, Axiom 4, and Koopmans' [9] argument, V can be expressed as

$$V(f) = \overline{W}(v(f_0), \overline{A}(\overline{G}(f_1), B({}_2f))),$$

for some continuous, nonconstant functions $\overline{W}$, $\overline{A}$, $\overline{G}$, and B. Now use axiom 8 to obtain, as shown by Koopmans (1960) [9], that A in (19) and B in (20) are homeomorphic and therefore B can be taken to equal A by a simple modification of the function $\overline{A}$. This leads to

$$V(f) = \hat{W}(v(f_0), \hat{A}(\hat{G}(f_1), A({}_2f))). \tag{20}$$

According to (19), for every $v(f_0)$, $\succ^*$ depends on ${}_1f$ only through $A({}_1f)$. Therefore, for all f_1,

$$A({}_1f) = \phi_1(a(f_1) + h({}_2f)),$$

for some strictly increasing and continuous function ϕ_1, where $h({}_2f) = b(f_2) + d({}_3f)$.

According to (20), for every $v(f_0)$ and $\hat{G}(f_1)$, $\succ^*$ depends on ${}_2f$ only through $A({}_2f)$. Therefore, for all ${}_2f$,

$$A({}_2f) = \phi_2(b(f_2) + d({}_3f)),$$

for some strictly increasing and continuous function ϕ_2.

According to (20), for every $v(f_0)$ and $A({}_2f)$, $\succ^*$ depends on f_1 only through $\hat{G}(f_1)$. Therefore,

$$a(f_1) \equiv G(f_1) = \phi_3(\hat{G}(f_1)),$$

for some strictly increasing and continuous function ϕ_3.

Now comparing the two previous equations for A implies that, for all f,

$$a(f_2) + h({}_3f) = \phi(b(f_2) + d({}_3f)),$$

where ϕ is some strictly increasing continuous function.

Lemma. *ϕ is affine.*

Proof. Let $x = f_2 \in \mathcal{X}$ and $y = {}_3f \in \mathcal{Y}$. We have

$$a(x) + h(y) = \phi(b(x) + d(y)),$$

where ϕ is increasing and continuous. Note that, since b, d, and U are continuous and nonconstant and X is connected, without loss of generality $I = \{b(x) + d(y) : x \in \mathcal{X},\ y \in \mathcal{Y}\}$ is a connected, nonempty interval. Choose $x_0 \in \mathcal{X}$ and $y_0 \in \mathcal{Y}$ arbitrarily, and define $\overline{a}(x) = a(x) - a(x_0)$, $\overline{h}(y) = h(y) - h(y_0)$, and $\overline{b}(\cdot)$ and $\overline{d}(\cdot)$ similarly. So

$$\begin{aligned}\overline{a}(x) + \overline{h}(y) &= \phi(\overline{b}(x) + \overline{d}(y) + b(x_0) + d(y_0)) - \phi(b(x_0) + d(y_0)) \\ &\equiv \overline{\phi}(\overline{b}(x) + \overline{d}(y)).\end{aligned}$$

Note that $\overline{\phi}$ is continuous on the connected nonempty interval $\overline{I} = I - b(x_0) - d(y_0)$, which contains 0, and that $\overline{a}(x) = \overline{\phi}(\overline{b}(x))$ and $\overline{h}(y) = \overline{\phi}(\overline{d}(y))$. So,

$$\overline{\phi}(\tilde{b} + \tilde{d}) = \overline{\phi}(\tilde{b}) + \overline{\phi}(\tilde{d}) \tag{21}$$

for all $\tilde{b} \in I_b = \{\overline{b}(x) : x \in \mathcal{X}\}$ and $\tilde{d} \in I_d = \{\overline{d}(y) : y \in \mathcal{Y}\}$.

Using (21), we can now show that $\overline{\phi}$ is linear, thus ϕ is affine. First, note that since $\overline{\phi}(0) = 0$, $\overline{\phi}(x) = -\overline{\phi}(-x)$; so we can focus on the positive part, $\overline{I}_+$, or the negative part, $\overline{I}_-$, of $\overline{I}$. Suppose, without loss of generality, that $\overline{I}_+ \neq \varnothing$. Consider any $b > b' > 0$ in $\overline{I}_+$ such that b and b' are rational. Then, by (21), $\overline{\phi}(b) = m\overline{\phi}(\frac{1}{n})$ and $\overline{\phi}(b') = m'\overline{\phi}(\frac{1}{n'})$ for m, m', n, $n' \in \mathbb{N}$. Since $\overline{\phi}(\frac{1}{n})n = \overline{\phi}(1) = \overline{\phi}(\frac{1}{n'})n'$, it follows that $\overline{\phi}(b) = \frac{b}{b'}\overline{\phi}(b')$. Since rationals are dense in $\overline{I}_+$ and $\overline{\phi}$ is continuous, $\overline{\phi}(b) = \frac{b}{b'}\overline{\phi}(b')$ holds for all $b, b' \in \overline{I}_+$, which implies linearity.

□

Since ϕ must be increasing, there exists $\alpha > 0$ such that $b(f_2) = \alpha a(f_2)$ and

$$d({}_3f) = \alpha h({}_3f) = \alpha(b(f_3) + d({}_4f)). \tag{22}$$

It follows that

$$\overline{V}(f) = \hat{u}(f_0) + G(f_1) + \alpha G(f_2) + d({}_3f).$$

Now restrict attention to streams f that are constant after $t = 3$—which correspond to consumption streams that are constant after $t = 3$. For such streams, $d({}_3f) = \hat{d}(f_3)$. By Axiom 8 and $\alpha > 0$,

$$G(f_3) \geq G(f_3') \iff \hat{d}(f_3) \geq \hat{d}(f_3').$$

So $\hat{d}(\cdot) = \varphi(G(\cdot))$ for some strictly increasing and continuous function φ. Again by Axiom 8,

$$\alpha G(f_2) + \varphi(G(f_3)) \geq \alpha G(f_2') + \varphi(G(f_3'))$$

if and only if

$$G(f_2) + \alpha G(f_3) + \varphi(G(f_3)) \geq G(f_2) + \alpha G(f_3) + \varphi(G(f_3)).$$

So

$$\alpha G(f_2) + \varphi(G(f_3)) = \alpha(G(f_2) + \alpha G(f_3) + \varphi(G(f_3))) + k,$$

which implies $\varphi(G)(1-\alpha) = \alpha^2 G + k$. Since φ must be strictly increasing, it follows that $\alpha < 1$. Finally, by iteratively applying (22) and relying on $\alpha \in (0,1)$, we get that $\succ^*$ can be represented by

$$V^*(f) = u(f_0) + \sum_{t=1}^{\infty} \alpha^t G(f_t).$$

Finally, it is easy to see that Axiom 9 holds if and only if u and G are increasing—using Lemma 2.

8.6 Proof of Theorem 4

By Debreu (1960) [3], $\succ$ can be represented by

$$w_0(c_0) + w_1(c_1) + w_2({}_2c),$$

for some continuous and nonconstant functions w_0, w_1, and w_2. By Theorem 3, $\succ$ is also represented by

$$u(c_0) + \alpha G(u(c_1) + g({}_2c)) + \alpha g({}_2c),$$

where $g({}_2c) = \sum_{t=2}^{\infty} \alpha^{t-1} G(U({}_tc))$. It follows that

$$u(c_0) + \alpha G(u(c_1) + g({}_2c)) + \alpha g({}_2c) = \xi \left[w_0(c_0) + w_1(c_1) + w_2({}_2c)\right] + \chi,$$

where $\xi > 0$ and $\chi \in \mathbb{R}$. This implies that

$$\alpha G(u(c_1) + g({}_2c)) + \alpha g({}_2c) = \xi \left[w_1(c_1) + w_2({}_2c)\right],$$

and therefore G must be affine.

References

[1] Andrew Caplin and John Leahy. Psychological expected utility theory and anticipatory feelings. *The Quarterly Journal of Economics*, 116(1):55–79, 2001.

[2] Gerard Debreu. Representation of a preference ordering by a numerical function. *Decision processes*, 3:159, 1954.

[3] Gerard Debreu. Topological methods in cardinal utility theory. *Mathematical methods in the social sciences*, 1959:16–26, 1960.

[4] Peter A Diamond. The evaluation of infinite utility streams. *Econometrica: Journal of the Econometric Society*, pages 170–177, 1965.

[5] Irving Fisher. The theory of interest. *New York*, 43, 1930.

[6] Shane Frederick, George Loewenstein, and Ted O'donoghue. Time discounting and time preference: A critical review. *Journal of economic literature*, 40(2):351–401, 2002.

[7] Faruk Gul and Wolfgang Pesendorfer. Temptation and self-control. *Econometrica*, 69(6):1403–1435, 2001.

[8] Takashi Hayashi. Quasi-stationary cardinal utility and present bias. *Journal of Economic Theory*, 112(2):343–352, 2003.

[9] Tjalling C Koopmans. Stationary ordinal utility and impatience. *Econometrica: Journal of the Econometric Society*, pages 287–309, 1960.

[10] Tjalling C Koopmans, Peter A Diamond, and Richard E Williamson. Stationary utility and time perspective. *Econometrica: Journal of the Econometric Society*, pages 82–100, 1964.

[11] Botond Kőszegi. Utility from anticipation and personal equilibrium. *Economic Theory*, 44(3):415–444, 2010.

[12] David Laibson. Golden eggs and hyperbolic discounting. *The Quarterly Journal of Economics*, 112(2):443–478, 1997.

[13] David Laibson, Andrea Repetto, and Jeremy Tobacman. Estimating discount functions with consumption choices over the lifecycle. Technical report, National Bureau of Economic Research, 2007.

[14] George Loewenstein. Anticipation and the valuation of delayed consumption. *The Economic Journal*, 97(387):666–684, 1987.

[15] George Loewenstein and Drazen Prelec. Anomalies in intertemporal choice: Evidence and an interpretation. *The Quarterly Journal of Economics*, 107(2):573–597, 1992.

[16] José Luis Montiel Olea and Tomasz Strzalecki. A simple axiomatization of quasi-hyperbolic discounting. Technical report, working paper, 2011.

[17] DavidG. Pearce. Nonpaternalistic sympathy and the inefficiency of consistent intertemporal plans. In MatthewO. Jackson and Andrew McLennan, editors, *Foundations in Microeconomic Theory*, pages 213–231. Springer Berlin Heidelberg, 2008.

[18] Edmund S Phelps and Robert A Pollak. On second-best national saving and game-equilibrium growth. *The Review of Economic Studies*, 35(2):185–199, 1968.

[19] Ariel Rubinstein. 'economics and psychology' the case of hyperbolic discounting*. *International Economic Review*, 44(4):1207–1216, 2003.

[20] Maria Saez-Marti and Jörgen W Weibull. Discounting and altruism to future decision-makers. *Journal of Economic theory*, 122(2):254–266, 2005.

[21] Paul A Samuelson. A note on measurement of utility. *The Review of Economic Studies*, 4(2):155–161, 1937.

[22] Marciano Siniscalchi. Dynamic choice under ambiguity. *Theoretical Economics*, 6(3):379–421, 2011.

[23] Robert Henry Strotz. Myopia and inconsistency in dynamic utility maximization. *The Review of Economic Studies*, 23(3):165–180, 1955.

[24] Eugen von Böhm-Bawerk. *Capital and interest: A critical history of economical theory*. Macmillan and Co., 1890.