Forward-Looking Behavior Revisited:
A Foundation of Time Inconsistency

Simone Galperti  Bruno Strulovici*
UC, San Diego  Northwestern University

May 19, 2014

Abstract

Regardless of its interpretation, the standard exponentially-discounted-utility model implies myopically forward-looking behavior: Preferences at a given period are completely determined by consumption in that period and preferences at the next period. This paper axiomatizes preferences that capture fully forward-looking behavior, delivering a new class of utility representations which includes quasi-hyperbolic ($\beta$-$\delta$) discounting as a particular case. These representations rationalize phenomena left unexplained not only by the standard model, but also by the $\beta$-$\delta$ discounting model. Time inconsistency in the form of present bias and other phenomena are necessary, logical consequences of fully forward-looking behavior and do not require time-varying preferences or psychological considerations absent from the standard model. The approach also delivers tractable, Bellman-type equations characterizing optimal consumption streams, as well as new insights for the welfare analysis of 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 Nageeb S. Ali, Nabil Al-Najjar, Eddie Dekel, Bart Lipman, Peter Klibanoff, Mark Machina, Paul Milgrom, David G. Pearce, Todd Sarver, Marciano Siniscalchi, and Asher Wolinsky, as well as seminar participants at UC San Diego, Stanford, Université de Montréal for useful comments. An early version of this project was developed by Strulovici and presented at Northwestern University and the 2011 SAET conference in Faro, Portugal under the title “Forward-Looking Behavior, Well-Being Representation, and Time Inconsistency.” Strulovici is grateful for financial support from the National Science Foundation and a Sloan Research Fellowship.
1 Introduction

Facing long-term decisions, forward-looking agents often satisfy the following two premises: (1) They take into account their preferences at many future dates, not just at the next instant; (2) They do not care about future consumption per se, but about how this consumption accords with their future preferences. For example, a young worker deciding on a retirement plan is motivated by his well-being upon retirement, not by how his decision today captures his preference for savings tomorrow, when he is still young. And this worker cares little about increasing some specific consumption upon retirement (say, traveling) unless such consumption is actually valued according to his preferences at that later time. Similarly, an undergraduate student choosing to pursue graduate studies cares about how he will feel about this choice throughout his career, not just about how he will feel about it while he is in graduate school.

While apparently innocuous, these premises are ruled out by the standard exponentially discounted utility (EDU) model. That model, just like any recursive model, requires that preferences today be entirely determined by consumption today and preferences at the next period only. This is a mathematical implication of those models, which holds regardless of any interpretation one might be tempted to give them: Today’s preferences cannot directly incorporate future preferences beyond the next period. This assumption is problematic, as illustrated by the previous examples. Marriage provides an even starker illustration of this: A young adult committing to get married in the next period (month or year) does not only care about how he or she will feel during that first period, but also about how marriage will reflect his or her preferences beyond that first period.

Samuelson proposed an interpretation of EDU in which preferences at any future period $t$ are completely captured by consumption at period $t$ or, in Samuelson’s words, “at every instant of time [the agent’s] satisfaction depends only upon the consumption at that time,” (Samuelson (1937), p. 159). This interpretation is unsatisfactory, however, as the agent must treat his future selves as myopic. It is also inconsistent, because the same EDU model applied to any future date $t$ stipulates that the agent will be forward-looking at that time. Another possibility is to simply give up on the interpretation that the agent cares about his future preferences. For example, the agent could be simply

---

1 Let $U_s = \sum_{t \geq s} \delta^{t-s} u(c_t)$ denote the agent’s total utility at time $s \geq 0$ from a consumption stream $(c_s, c_{s+1}, \ldots)$ in the standard model. We have $U_0 = u(c_0) + \delta U_1$. The total utility of the agent at time 0, which represents his preferences at that time, cannot directly depend on $U_s$ for $s \geq 2$.

2 Samuelson himself was uncomfortable with the assumption that $u(c_t)$ represents the agent’s ‘satisfaction’ at time $t$. He also viewed it as “completely arbitrary to assume that [the agent] behaves so as to maximize” the exponentially discounted sum of instantaneous utilities, $\{u(c_t)\}_{t \geq 0}$.

2
planning future consumption irrespective of his future preferences. As argued earlier, however, this alternative interpretation is equally problematic, as it does not provide an explicit and coherent link between future consumption and future preferences.

This paper proposes a theory of ‘fully’ forward-looking behavior grounded on the above two premises, which sheds light on some well-documented behaviors. In particular, it provides a novel explanation for time inconsistency in the form of present bias.\(^3\) It turns out that present bias (not just any form of time inconsistency) must arise whenever the agent takes into account all future preferences, regardless of any functional assumption. Assuming fully forward-looking behavior, present bias is completely unrelated to any notion of ‘irrationality,’ myopia, or ‘excessive’ taste for the present.

Another contribution is to introduce and characterize, through a simple axiomatization, a new, tractable class of intertemporal utility models for fully forward-looking preferences over infinite consumption streams. Preferences in this class can be represented by

\[
U(c_0, c_1, \ldots) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U(c_t, c_{t+1}, \ldots)),
\]

(1)

where \(\alpha\) and \(G\) may be interpreted as the discount factor for future preferences and ‘utility’ (at time zero) from future well-being.\(^4\) The parameter \(\alpha\) and function \(G\) must satisfy joint restrictions coming from the axioms. In particular, \(G\) must be \((1 - \alpha)/\alpha\) Lipschitz continuous. This property implies that the agent exhibits impatience, a predilection for receiving better consumption sooner than later. It is natural to assume that \(G\) is concave. Intuitively, the agent exhibits decreasing marginal utility in future well-being. Given \(u, \alpha,\) and \(G,\) with \(G\) satisfying the Lipschitz restriction, there exists a unique solution \(U\) to (1) within the class of ‘\(H\)-continuous’ functions.\(^5\)

Fully forward-looking behavior provides a clear explanation for consumption dependence, the frequently observed phenomenon that intertemporal trade-offs between consumption at any two given periods depends on consumption at other periods. Specifically, the paper explains why, and how, the effective discount rate between two periods depends on consumption at other periods. To illustrate this question, consider the following two examples: (1) A worker learns that his firm will have to shut down within a year, and

\(^3\)Intuitively, an agent exhibits present bias if there are situations in which he is willing to give up, at time 0, an opportunity of consumption at some future time \(t\) in exchange for better consumption at \(t + 1\), but not willing to give up the same good consumption at 0 in exchange for the same better consumption at 1.

\(^4\)This is just a convenient of interpretation of how future preferences affect today’s preference.

\(^5\)\(H\)-continuity roughly means that the agent’s total utility is almost the same for streams that are identical over a long enough horizon. We introduce that concept to deal with infinite horizon problems and show that preferences must satisfy this strong notion of continuity, despite assuming a much weaker axiom of continuity at the outset.
must adapt his consumption choices to this news; (2) A minor learns that he will receive a large bequest that will be available upon reaching adulthood. How do the economic events described in these examples affect an agent’s intertemporal preferences? Suppose that $G$ is concave. In example (1), the theory predicts that the worker will exhibit more patience—i.e., less discounting—after learning about his future job loss. Intuitively, his future well-being has gone down as the result of the news, and by concavity of $G$, his marginal return for higher well-being, and hence higher future utility from consumption, has gone up. This makes him more patient. Similarly, in example (2), the minor can exhibit less patience—i.e., more discounting—after learning about his future large bequest.\footnote{For evidence consistent with these behaviors, see Frederick et al. (2002).}

This ‘consumption dependence’ cannot arise in any intertemporal model that is separable in consumption at different periods, such as the EDU model and the quasi-hyperbolic ($\beta$-$\delta$) model. More generally, the paper derives an explicit formula for the agent’s discount factor between 0 and $t$. For general $G$, this factor will depend (through well-being) not only on consumption at $t$—as empirically observed (see Frederick et al. (2002))—but also on the entire infinite consumption stream—as suggested by Fisher (1930).

The theory also offers a new axiomatization of $\beta$-$\delta$ discounting as a special case of representation (1) when $G$ is linear.\footnote{Section 3 discusses previous axiomatizations of $\beta$-$\delta$ discounting (Hayashi (2003); Olea and Strzalecki (2013)).} Far from being ‘irrational,’ a $\beta$-$\delta$ agent in this theory treats immediate consumption (i.e., $c_0$) and future preferences (i.e., $U(c_t, c_{t+1}, \ldots)$) as different entities; he takes into account his preferences in all future periods in a time separable and stationary way (as defined below); and he trades off consumption at different periods in a way that does not depend on $U(c_t, c_{t+1}, \ldots)$ at other periods—this will correspond to $G$ being linear in (1).

Finally, by constructing a rigorous framework in which present bias is not assumed but is derived from a more primitive property of preferences, the paper also provides a clean way to do welfare analysis with time-inconsistent agents. The choice of a welfare criterion depends on how one interprets fully forward-looking preferences. One possibility is that the agent is simply planning his future consumption and takes into account his future preferences. According to this ‘planning’ interpretation, ex ante the agent does not experience any utility from imagining his future well-being, but is instead already performing the job of a social planner. This interpretation suggests a libertarian approach to welfare analysis: to use the agent’s ex-ante preference as a measure of welfare.

The other possibility is that the agent cares about his future preferences because they directly affect his current well-being. In this case, the agent experiences anticipatory
utility of a specific nature: He does not get such utility from the act itself of imagining future consumption *per se*, but from anticipating how this consumption will affect his future well-being.\(^8\) This second interpretation suggests a welfare criterion that aggregates the agent’s well-being at all periods. In the case of agents with \(\beta-\delta\) preferences, this interpretation turns out to give a new, rigorous justification—based on representation (1)—for the criterion that has been often assumed in the past: the EDU formula resulting from setting \(\beta = 1\). This criterion, of course, is *time consistent* and therefore is conducive to a more tractable welfare analysis. By contrast, in the case of EDU agents, a similar anticipatory interpretation would lead to a time-inconsistent planner.\(^9\)

Despite their versatility, the models obtained by representation (1) are quite tractable. Section 6.2 considers, as an illustration, the problem of an agent who has to choose among consumption streams subject to some budget constraint. It shows that we can solve this problem by solving a maximization that is similar to a standard Bellman equation and is guaranteed to have a unique solution.

To avoid confusion, this paper is purely concerned with preferences, not with the dynamic-decision conflicts that can arise from time inconsistency, or the way such conflicts might be resolved. These issues are entirely distinct from this paper’s subject, and much confusion may arise by considering them together. The paper, for instance, draws a clear distinction between invariant, time-inconsistent preferences and choice reversals. Their connections, however, are briefly discussed in Section 10.

### 2 Axiomatization and Present Bias

This section briefly introduces the axiomatization leading to representation (1) and explains why fully forward-looking agents necessarily exhibit present bias.

#### 2.1 ‘Well-Being’ Representation

The axiomatization, starts by providing a general representation of fully forward-looking preferences, given by

\[
U(c) = V(c_0, U_1, U_2, \ldots),
\]

where \(U_t = U(c_t, c_{t+1}, \ldots)\) represents the agent’s preference at \(t\). As noted, we will sometimes interpret \(U_t\) as the agent’s *well-being* in period \(t\). This interpretation, while

\(^8\)Section 3 reviews the literature on anticipatory utility of the first kind, explaining how it differs from this paper.

somewhat restrictive, is helpful to convey the intuition for this paper’s results.\textsuperscript{10} Representation (2) highlights two conceptual aspects of the theory. Firstly, future well-being, not future instantaneous utility, determines today’s well-being (together with today’s consumption). Secondly, today’s consumption is inherently different from (future) well-being, a more general and comprehensive entity. Such a distinction links this paper with Böhm-Bawerk’s (1890) and Fisher’s (1930) early work on intertemporal choice. The general representation (2) hinges on the following simple axiom. Given two streams with equal initial consumption, if the agent is indifferent between their continuations starting at any future date (i.e., $c = (c_t, c_{t+1}, \ldots)$), then he cannot prefer either stream—for example, because they allocate consumption differently over time. This ensures that the agent’s current preference depends on future consumption only through future preferences.

2.2 Present Bias: Intuition

Despite preferences being invariant over time—a hypothesis maintained throughout—fully forward-looking behavior implies a specific form time inconsistency: present bias. This result may seem paradoxical, for a fully forward-looking agent cares more about his future preferences than in standard models. This behavior has a simple explanation which, while resolving the paradox, differs significantly from one’s usual understanding of present bias. To understand the source of present bias, suppose that a fully forward-looking agent is indifferent, at time 0, between two continuation streams $c$ and $c'$ that start at time 1, with $c_1 > c'_1$, $c_2 < c'_2$ and $c_t = c'_t$ for $t > 2$. Suppose also that, everything else equal, the agent likes higher consumption better. So $c'$ must provide a higher well-being than $c$ at time 2. Since this higher well-being at 2 is directly taken into account at 0, the agent can only be indifferent at 0 if $c$ provides a higher well-being than $c'$ at time 1. Finally, preferences are time invariant, which implies that if the agent had to choose between $c$ and $c'$ starting at time 0, he would strictly prefer $c$, which has higher immediate consumption, and hence exhibits present bias.

2.3 Axiomatization of the ‘G’-Representation (1)

To obtain a more tractable model of fully forward-looking behavior, we explore the consequences of imposing time separability and stationary dependence on future preferences/well-being.
being, in contrast to future consumption/instantaneous utility. These additional axioms lead to the new class of utility representations (1) Separability in immediate consumption and future well-being follows from an adaptation of Debreu’s (1960) separability axiom. Regarding stationarity, in general, it says that ex ante the agent does not change how he ranks options, if all options are simply 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, postponing consumption streams that start at a future \( t \) 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. Therefore, this paper’s stationarity axiom refers only to future well-being; it also differs substantively from Koopmans’s (1960) stationarity as well as from Olea and Strzalecki’s (2013) quasi-stationarity (see Sections 3 and 9).

3 Literature Review

Several empirical studies have criticized EDU, and several papers have addressed its flaws in various ways,\(^ {11}\) For instance, Loewenstein (1987), Caplin and Leahy (2001), and Kőszegi (2010) add to EDU a dimension called anticipation utility. That is, besides the usual instantaneous utility at date \( t \), consumption at \( t \) also generates a distinct utility at all \( t' < t \). Such a utility captures the effects of feelings, like excitement or anxiety, that the agent harbors today when thinking about future events. This is different from anticipating, in the sense of imagining, the future feelings arising when those events occur.\(^ {12}\)

Therefore, anticipation utility differs from fully forward-looking behavior. To see this, consider a simple two-period example. At time 2, the agent consumes a concert and, while there, a drink. His time-2 utility is \( u(s, d) \), which is strictly increasing in the seat’s proximity to the stage, \( s \), and the drink’s quality, \( d \). At time 1, the agent derives anticipation utility in the form of excitement from thinking about being at the concert, but no anticipation utility from the drink. Let the agent’s excitement utility be the strictly increasing function \( e(s) \). At 1, the agent has $100 that he can use to buy a concert ticket; whatever is left will go for the drink. At 1, he maximizes \( e(s) + u(s, d) \),

\(^{11}\) See also Epstein (1983), Epstein and Hynes (1983).

\(^{12}\) In Böhm-Bawerk’s (1890) words, “We must distinguish between two fundamentally distinct things […]. It is one thing to represent to ourselves, or imagine, a future pleasure or future pain […]. It is quite another thing to experience, in this imagination itself, […] an actual present pleasure of anticipation.” (Book V, Ch. 1, p. 239).
whereas at 2 he would maximize $u(s, d)$. So, for any $d$, his marginal utility from $s$ is higher at 1. He is therefore willing to spend more for a ticket at 1 than at 2; i.e., he is time inconsistent. By contrast, with only two periods, a fully forward-looking agent maximizes $u(s, d)$ in both periods and consequently is time consistent.

Although conceptually different, anticipation-utility models can lead to time inconsistency and accommodate some anomalies of EDU addressed in the present paper. This paper, however, does so using the notion of forward-looking behavior and therefore complements explanations based on anticipation utility.

Phelps and Pollak (1968) introduced $\beta$-$\delta$ 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 itself than about any future generation—composed, after all, of unborn strangers. In this view, the $\beta$-$\delta$ formula, though simple, has a natural justification. Laibson (1997) makes a significant conceptual leap, by applying the same formula to individual decision-making. He justifies it based on its ability “to capture the qualitative properties” of hyperbolic discounting, which has received substantial empirical support. Nonetheless, with a single agent, 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 $\beta$-$\delta$ model in settings with a single agent (see, e.g., Hayashi (2003); Olea and Strzalecki (2013)). These axiomatizations differ from that of the present paper as follows. First, they continue to view the agent as caring about instantaneous utilities. Within this framework, they replace Koopmans’s (1960) stationarity—clearly violated by $\beta$-$\delta$ discounting—with quasi-stationarity, namely stationarity from the second period onward. Assuming quasi-stationarity, however, does not address the conceptual issue posed by Laibson’s model. Second, to obtain the $\beta$-$\delta$ representation, they need to ensure that current and future instantaneous utilities are cardinally equivalent. Olea and Strzalecki’s axioms permit useful experiments to identify and measure $\beta$ and $\delta$, but they seem difficult to interpret. By contrast, the present paper views the agent as caring about immediate consumption and future well-being (i.e., preferences). As a result, its stationarity notion—involving only well-being—has a more natural interpretation and can offer a rationale for the $\beta$-$\delta$ formula in Laibson’s setting. The paper also shows that this formula is tightly linked with an intuitive property: intertemporal consumption trade-offs do not depend on intermediate and future well-being.

As noted, this paper is related to the literature on intergenerational altruism at a super-
ficial, formal level, but it differs at a deeper, conceptual level. That literature considers a collection of individuals, not a single agent. Clearly, an agent’s time preference and altruism towards other individuals are distinct aspects of behavior. More specifically, Saez-Martí and Weibull (2005) derive the mathematical equivalence between the $\beta$-$\delta$ formula and expression (1) with linear $G$. But they do not provide an axiomatic foundation of either representation and, hence, do not address the conceptual issues mentioned before. Pearce's (2008) work on nonpaternalistic sympathy considers a cake-eating model with finitely many generations. It assumes that each generation’s well-being depends on its consumption as well as the well-being of all other or only future generations, and focuses on equilibrium analysis, proving general inefficiency results. In a similar vein, Bergstrom (1999) studies the systems of utility functions that include altruism towards others, focusing on the infinite regress that it may generate. Finally, in his study of hedonistic altruism and welfare, Ray (2014) examines welfare criteria that aggregate the well-being of altruistic, time-consistent, generations and that are formally similar to those in Section 7.\(^\text{13}\)

Finally, a literature in philosophy studies the normative question of how an individual’s different ‘selves,’ corresponding to different times, should be weighted in a given decision problem (Parfit [1971, 1976, 1982]). Similarly, economists have studied normative criteria for intergenerational problems (see, e.g., for catastrophe risk). These normative issues, while interesting, are orthogonal to the question considered here.

4 Preliminaries

This paper considers how an agent chooses consumption streams. At each time $t$, the set of feasible consumption levels is $X$, a connected, separable, metric space. Time is discrete with infinite horizon: $T = \{0, 1, 2, \ldots\}$. For $t \geq 0$, the set of consumption streams starting at $t$ is $tC \subseteq X^T$—elements of this set will be denoted by $tC = (c_t, c_{t+1}, \ldots)$. For each $t$, the agent has a preference relation $\succ^t$ with domain $tC$.\(^\text{14}\) In principle, $\succ^t$ and $tC$ could vary over time. In this paper, however, they are time invariant.

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

\(^{13}\)Jackson and Yariv (2011) study the problem of how to aggregate preferences of heterogeneous, time-consistent individuals in collective dynamic decisions and show that natural procedures lead to time inconsistency.

\(^{14}\)This paper continues to assume that, at each $t$, preferences do not depend on past consumption. Relaxing this assumption is beyond the paper’s scope.
The set $C$ is endowed with the sup-norm: $\|c - c'\|_C = \sup_t d(c_t, c'_t)$ where $d$ is the metric on $X$.

**Remark 1.** For any $c, c' \in C$, the expression $c \succ^t c'$ means the following. At time $t$ the agent always chooses $c$ over $c'$, if at that time he has to commit to either stream for the entire future. This commitment assumption is essential for interpreting the agent’s choice as his ranking of streams. Otherwise, at $t$ the agent may choose $c$ because he likes $c_t$ better than $c'_t$ and anticipates that in the future he will abandon $c$ for some other stream. In this case, at $t$ the agent is not actually comparing $c$ and $c'$. Hence, his choice does not tell us anything about his preference over $c$ and $c'$. Moreover, without the commitment assumption, we cannot meaningfully examine whether the agent’s preferences are time consistent.

A basic premise of this paper is that $\succ$ has a utility representation. As shown in Appendix B, this property follows from standard axioms. To simplify exposition, it is stated here as an assumption.

**Assumption 2 (Utility Representation).** There is a continuous function $U : C \to \mathbb{R}$ such that $c \succ c'$ if and only if $U(c) > U(c')$. Moreover, $U$ is nonconstant in the first and some other argument.

**Remark 2.** In the rest of the paper, $U(\cdot, c)$ will be interpreted as the agent’s well-being (i.e., total utility) generated by consumption stream $\cdot, c$. This is done only for the sake of conveying intuitions more easily.

Since this paper is interested in forward-looking behavior, by assumption, well-being at $t$ depends on consumption after $t$. It is also natural that well-being depends on immediate consumption. For future reference, let $U$ be the range of $U$. Note that $U$ is an interval because $U$ is continuous and non-constant and $X$ is connected. Since preferences are time invariant (Assumption 1), for simplicity, hereafter we take the perspective of $t = 0$.

## 5 Forward-looking Behavior and Time Inconsistency

### 5.1 ‘Well-Being’ Preference Representation

This paper considers an agent who, at each $t$, is forward-looking in the sense that he cares about his future preferences, represented by well-being $U$. That is, it focuses on $\succ$ that has the following recursive representation.
Definition 1 (Well-Being Representation). Preference $\succ$ has a well-being representation if and only if

$$U(c) = V(c_0, U_1(c), U_2(c), \ldots)$$

for some function $V : X \times \mathbb{R}^N \to \mathbb{R}$ that is nonconstant in $c_0$ and $U(t c)$ for some $t > 0$.\(^{15}\)

So, how the agent ranks streams $c$ and $c'$ at 0 depends only on the immediate consumption levels $c_0$ and $c'_0$, and on how he ranks continuation streams $t c$ and $t c'$ for at least some future period $t$. In other words, the well-being from any $c$ depends only on immediate consumption $c_0$ and, for at least some future $t$, on per-period well-being from $t c$. Well-being is therefore conceptually different from the immediate utility from a single consumption event: Well-being captures an aggregate of sensorial pleasure from immediate consumption as well as purely mental satisfaction (or dissatisfaction) from future well-being. While such an aggregation seems difficult, it is somehow performed by any forward-looking agent who must choose current and future consumption. The definition is recursive, as one should expect: For a forward-looking agent, well-being today involves well-being in the future. Finally, note that, at this stage, $V$ may be strictly decreasing in $U(t c)$ for some $t > 0$.

Axiom 1 is the key to obtaining representation (3). It captures the idea that streams starting tomorrow affect well-being today, only through the well-being that they generate at each future time.

Axiom 1. If $t c \sim t c'$ for all $t > 0$, then $(c_0, t c) \sim (c_0, t c')$.

Axiom 1 rules out the possibility that the agent prefers $c$ over $c'$ because, despite generating the same stream of immediate consumption and future well-being, they allocate future consumption differently over time.

Theorem 1. Axiom 1 holds if and only if $\succ$ has a well-being representation.

Proof. Let $f_0(c) = c_0$ and, for $t > 0$, $f_t(c) = U(t c)$. Also, let $f(c) = (f_0, f_1, f_2, \ldots)$ and

$$\mathcal{F} = \{f(c) : c \in C\}.$$  \(^{(4)}\)

($\Rightarrow$) First, define equivalence classes on $C$ as follows: $c$ is equivalent to $c'$ if $f_t(c) = f_t(c')$ for all $t \geq 0$.\(^{16}\) Let $C^*$ be the set of equivalence classes of $C$, and let the function $U^*$ be

\(^{15}\)Of course, one could allow preferences to differ across time, yet obtain for each $\succ^t$ a well-being representation. In this case, $U$ and $V$ will depend on $t$.

\(^{16}\)In general, there may be several consumption streams in an equivalence class. For example, suppose that $U(c) = c_0 + c_1 + c_2 + c_3$, and let $c = (1, 1, -1, -1, 1, 1, -1, -1, \ldots)$ and $c' = (1, -1, -1, 1, 1, -1, -1, 1, 1, \ldots)$.\(^{11}\)
defined by $U$ on $C^*$. Then, the function $f^* : C^* \to \mathcal{F}$, 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 1, $V$ is a well-defined function, and $V(f(c)) = U(c)$ for every $c$. By Assumption 2, $V$ is nonconstant in $f_0$ and $f_t$ for some $t > 0$.

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

Note that Axiom 1 is weak: It requires that the agent be indifferent at 0 between two streams, only if he is indifferent between their truncations at all future dates, not just next one. Clearly, if at 0 the agent cares only about his well-being at 1—as under EDU—Axiom 1 holds. As clarified in Section 9, by allowing current preference to depend on future preference in a richer way, Axiom 1 represents a key departure of this paper from the set of axioms characterizing EDU (Koopmans [1960, 1964]).

### 5.2 Time (In)consistency

Among the properties of intertemporal preferences, time consistency is perhaps the most prominent and studied one. Therefore, as a preliminary step, we consider the relation between time consistency and the extent to which the agent is forward-looking—that is, how $V$ depends on future well-being. Surprisingly, there is a tension between being forward-looking and having time-consistent preferences.

We can capture the idea that preferences are consistent over time as follows (see, e.g., Siniscalchi (2011)). Recall that $\succ^t$ is the agent’s preference at time $t$.

**Definition 2 (Time Consistency).** If $1c \sim^1 1c'$, then $(c_0, 1c) \sim^0 (c_0, 1c')$. If $1c \succ^1 1c'$, then $(c_0, 1c) \succ^0 (c_0, 1c')$.\(^\text{17}\)

To understand this definition, suppose that at time 1 the agent will have the opportunity to buy or not a ticket to go to Hawaii at $t > 1$. Does he rank these two options in the same way at 1 and at 0? Recall that these rankings are revealed by the agent’s choices at 1 and at 0—which, by Assumption 1, are generated by the same preference. Definition 2

\(^\text{17}\)Definition 2 looks similar to the stationarity axiom in Koopmans [1960, 1964]. However, time consistency and stationarity are conceptually very different (see Section 9).
first requires that, if at 1 the agent is indifferent between buying the ticket or not, then at 0 he should also be indifferent between buying or not at 1. Similarly, if at 1 the agent prefers not to buy the ticket, then at 0 he should also prefer not to buy it at 1.

Proposition 1 shows that an agent who cares about his well-being beyond the immediate future cannot be time consistent. The purpose of this preliminary result is simply to identify and highlight a possible source of time inconsistency. This source corresponds to an intuitive, more general form of forward-looking behavior than in the EDU model: a preference that in each period depends on well-being beyond the subsequent period.

**Proposition 1.** Preference $\succ$ satisfies time consistency if and only if $V(c_0, U(1c), U(2c), \ldots) = V(c_0, U(1c))$, for all $c \in C$, and $V$ is strictly increasing in its second argument.

It is common to view time consistency of preferences as the norm and time inconsistency as an exception. Proposition 1 reverses this view. If we deem natural that an agent should care about his well-being beyond the immediate future, then we have to conclude that time inconsistency should be the norm, rather than the exception.

To see the intuition for Proposition 1, suppose that $1c$ corresponds to buying a ticket to go to Hawaii at $t > 1$, and that, at 1, the agent slightly prefers not to buy the ticket. Now imagine that, at 0, he can choose whether to buy the ticket at 1. If at 0 the agent cares directly about his well-being beyond 1—i.e., when he will be at Hawaii—he should then strictly prefer to buy the ticket at 1. Intuitively, from the perspective of 1, the negative effect on immediate well-being of spending money for the ticket just offsets the positive effect of higher period-$t$ well-being. But at 0, since the agent cares directly about well-being beyond 1, he weighs more the positive effect of the Hawaii trip, thus strictly preferring to buy the ticket at 1. More generally, this mechanism can lead to situations in which, at 1, the agent wants to revise plans made at 0.

As noted, EDU satisfies

$$U(c) = u(c_0) + \delta U(1c) = V(c_0, U(1c)).$$

When the preference at $t$ depends only on immediate consumption and the preference at $t + 1$, we shall call the agent ‘myopically forward-looking.’ On the other hand, we

---

18 One could consider a weaker version of the second part of Definition 2: If $1c \succ 1c'$, then $(c_0, 1c) \succ^0 (c_0, 1c')$. This version, however, would be at odds with time consistency, unless the agent is fully myopic.

19 All omitted proofs are in Appendix B.

20 The point of Proposition 1 should not be misunderstood. It is well known that a preference is time consistent if and only if, at each $t$, it has a specific recursive representation in which $U(c)$ depends only on immediate consumption, $c_0$, and continuation utility $U(1c)$.

21 Another interpretation is that such an agent fails to realize that, in the future, he will continue to
shall call the agent ‘fully forward-looking’ if, at all $t$, his preference depends directly on his preferences at all $s > t$.\footnote{Of course, one could also consider the case in which $V$ depends on $U_{(t)}c$ up to some finite $t > 1$.}

### 5.3 Present Bias

In general, time inconsistency can take different forms. So a natural question is whether fully forward-looking behavior implies a specific form. We already know from Assumption 1 that, here, time inconsistency cannot arise because preferences change over time. For example, it cannot take the form that at 0 the agent prefers higher consumption for every period, but at 1 he prefers lower consumption for every period. Indeed, fully forward-looking behavior implies a form of time inconsistency that involves changes in intertemporal trade-offs depending on whether these occur in the present or in the future. In short, it implies present-bias.

**Definition 3 (Present Bias).** Let $x, y, w, h \in X$ and $c' \in C$. Suppose $(x, c) \succ (y, c)$ for all $c$ and $(z_0, \ldots, z_{t-1}, x, w, c') \sim (z_0, \ldots, z_{t-1}, y, h, c')$ for some $t > 0$, then $(x, w, c') \succ (y, h, c')$.

Intuitively, this definition says the following. Suppose that, fixing consumption in all other periods, the agent always strictly prefers consumption $x$ to $y$. Also, suppose there is a consumption $h$ (e.g., a trip to Hawaii) that can make him indifferent at 0 between getting $x$ or $y$ at some future $t$, provided $y$ is ‘compensated’ with $h$ at $t+1$. Then, given the same choice at 0, a present-biased agent continues to strictly prefers $x$ to $y$, thus pursuing immediate gratification.

**Proposition 2.** If $V(c_0, U_{(1)c}, U_{(2)c}, \ldots)$ is strictly increasing in $U_{(t)c}$ for all $t > 0$,\footnote{Section 6.1 will introduce an axiom that captures this property (Axiom 5).} then it implies present bias.

As shown in the proof of Proposition 2, its conclusion continues to hold if in Definition 3 both $w$ and $h$ occur at some periods $s$ after $t+1$.

Though perhaps surprising, this result has a simple intuition. Set $t = 1$ in Definition 3. To offset the better time-1 prospect of the stream with $x$ and make the agent indifferent at 0, the stream with $y$ at 1 must provide more future well-being. Thus, this stream relies more on future well-being to induce indifference at 0. As the streams are brought forward in time, they both lose some of their future well-being value. But this effect\footnote{care about future well-being.}
penalizes more the stream with \( y \), which relied more on future well-being. As a result, the stream with \( x \) is now strictly preferred.

6 Refining the Well-being Representation

6.1 Time Separability and Well-being Stationarity

This section refines the general representation \( V \) in (3), by considering preferences that satisfy some form of time separability and stationarity. These properties will deliver further implications on behavior and tractability.

The first two axioms imply that \( \succ \) is time separable—that is, separable in immediate consumption and future well-being, as well as across future well-being. Intuitively, this property means that future well-being does not affect how the agent enjoys immediate consumption, nor does well-being at \( t \) affect how he enjoys at time 0 the well-being at another time \( s \). Though plausible, these effects may work in different directions. Favoring anyone seems, at this stage, arbitrary. Hence, the axioms rule out these effects. To state Axiom 2, let \( \Pi \) consist of all unions of subsets of \( \{\{1\}, \{2\}, \{3, 4, \ldots\}\} \).

**Axiom 2 (Immediate-Consumption and Well-Being Separability).** Fix any \( \pi \in \Pi \). If \( c, \hat{c}, c', \hat{c}' \in C \) satisfy

(i) \( t_c \sim t_{\hat{c}} \) and \( t_{c'} \sim t_{\hat{c}'} \) for all \( t \in \pi \),
(ii) \( t_c \sim t_{c'} \) and \( t_{\hat{c}} \sim t_{\hat{c}'} \) for all \( t \in T \setminus \pi \),
(iii) 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 \( \hat{c} \succ \hat{c}' \).

To further understand Axiom 2, consider \( \pi = \{1\} \), for example. Suppose that \( c \) and \( c' \) yield the same immediate consumption and well-being in all periods except at 1, and that the agent prefers \( c \) to \( c' \). If for both \( c \) and \( c' \) we do not change well-being at 1 but change immediate consumption and well-being in all other periods in the same way, thus obtaining \( \hat{c} \) and \( \hat{c}' \), then the agent should prefer \( \hat{c} \) to \( \hat{c}' \). Axiom 2 is inspired by Debreu’s (1960) and Koopmans’s (1960) separability axioms. It differs to the extent that it requires that certain consumption streams be indifferent, rather than that certain consumption events be equal. This is because we want separability in well-being, which can be the same across streams allocating consumption differently over time.

Axiom 3 is of technical nature; it ensures that \( \succ \) does depend on well-being at time 1, 2, and 3 (Debreu’s essentiality condition).
Axiom 3 (Essentiality). There are \( x, x', y, y' \in X \) and \( c, c' \in C \) such that \( (z, x, \hat{c}) \succ (z, x', \hat{c}) \), \( (z', z'', y, c'') \succ (z', z'', y', c'') \), and \( (w, w', w'', c) \succ (w, w', w'', c') \) for some \( z, z', z'', w, w', w'' \in X \), and \( \hat{c}, c'' \in C \).

Axiom 4 captures the idea that \( \succ \) depends on future well-being in a stationary way. Intuitively, stationarity means that, at time 0, the agent does not change how he ranks consumption events, simply because they are moved to a subsequent date. In the general representation \( V \), however, how the agent trades off at 0 well-being at, say, \( t \) and \( t + 1 \) can depend on what \( t \) is. But this non-stationarity limits the model’s applicability across different settings. Of course, requiring stationarity is reasonable only if postponing a consumption event does not change its nature. Since in this paper instantaneous consumption and future well-being are conceptually different, Axiom 4 requires stationarity only with respect to future well-being.

Axiom 4 (Well-Being Stationarity). If \( c, c' \in C \) satisfy \( c_0 = c'_0 \) and \( c \sim_1 c' \), then
\[
(c_0, 2c) \succ (c'_0, 2c') \iff c \succ c'.
\]

Intuitively, Axiom 4 says the following. Suppose \( c \) and \( c' \) differ only in well-being from period 2 onward. Then, if we drop consumption at 1 and shift both continuation streams back one period, the agent should rank the new streams as he ranked \( c \) and \( c' \). Section 9 compares Axiom 4 with Koopmans’s (1960) stationarity and Olea and Strzalecki’s (2013) quasi-stationarity.

Finally, Axiom 5 captures the natural case in which a fully forward-looking agent is better off at 0 when his future well-being improves.

Axiom 5 (Well-being Monotonicity). If \( c, c' \in C \) satisfy \( c_0 = c'_0 \) and \( c \succ_1 c' \) for all \( t > 0 \), then \( c \succ c' \). Moreover, if \( c \succ_1 c' \) for some \( t > 0 \), then \( c \succ c' \).

Theorem 2 (Additive Well-Being Representation). Axioms 1-4 hold if and only if the function \( U \) may be chosen so that
\[
U(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U(tc)).
\]
where \( u : X \to \mathbb{R} \) and \( G : U \to \mathbb{R} \) are continuous, nonconstant functions and \( \alpha \in (0, 1) \). Moreover, if \( \hat{U} \), \( \hat{u} \), \( \hat{c} \), and \( \hat{G} \) represent the same \( \succ \) as in (5), then \( \hat{\alpha} = \alpha \), and there exist \( a > 0 \) and \( b \in \mathbb{R} \) such that \( \hat{U}(c) = aU(c) + b \), \( \hat{u}(x) = au(x) + b \), and \( \hat{G}(\hat{U}) = aG(U - b/a) \) for all \( c, x, \) and \( \hat{U} \). Finally, Axiom 5 holds if and only if \( G \) is strictly increasing.

According to Theorem 2, the agent derives an instantaneous utility \( u \) from immediate
consumption, as usual, and a per-period utility $G$ from future well-being. Moreover, he discounts utility from future well-being exponentially.

One might wonder whether expression (5) is always well-defined for any $G$. The effects of future well-being on preceding well-being may inevitably lead the summation in (5) to diverge to infinity. However, by Assumption 2—i.e., by Axioms 7-8 in Appendix 12.1—$U$ is a nonconstant representation of $\succ$ with values in the interval $U \subset \mathbb{R}$. Therefore, there always exist streams $c$ such that $U(c)$ is bounded. This implies joint restrictions on $\alpha$ and $G$.

In the natural case of increasing $G$, Proposition 3 identifies a sufficient (and almost necessary) condition for (5) to be well-defined, as well as other properties of $G$. It also shows that the function $U$ in (5) is such that the effect on current well-being of changes in future consumption becomes arbitrarily small, if such changes occur sufficiently far in the future.

**Definition 4.** A function $U : C \to \mathbb{R}$ is $H$-continuous if, for every $\varepsilon > 0$, there exists a time $T(\varepsilon)$ such that the following holds: If $c, \tilde{c} \in C$ satisfy $c_t = \tilde{c}_t$ for $t \leq T(\varepsilon)$, then $|U(c) - U(\tilde{c})| < \varepsilon$.

**Proposition 3.**

(i) Under axioms 1-5, in representation (5) $U$ is $H$-continuous, $G$ is bounded, and for $\nu', \nu \in U$

$$\frac{1 - \alpha}{\alpha} |\nu' - \nu| > |G(\nu') - G(\nu)|.$$  

(ii) Suppose $G$ is strictly increasing, bounded, and $K$-Lipschitz continuous with $K < \frac{1 - \alpha}{\alpha}$, i.e., for all $\nu', \nu \in U$

$$K |\nu' - \nu| \geq |G(\nu') - G(\nu)|.$$  

Then, there is a unique $H$-continuous function $U : C \to \mathbb{R}$ that solves (5).

This result helps to choose $G$ appropriately in applications. Note that boundedness of $G$ implies that future well-being can have only a limited impact on current well-being. Intuitively, the agent cannot become infinitely happy or unhappy just from imagining his future well-being. No bound applies, however, to the immediate instantaneous utility.

Proposition 3 also implies that a fully forward-looking agent is always impatient—even if his future well-being at each date depends on subsequent well-being and he correctly anticipates this.

**Definition 5** (Impatience). Let $x, y \in X$ be such that $(x, c) \succ (y, c)$ for all $c \in C$. Preference $\succ$ exhibits impatience if, for any $t > 0$, $c^x \succ c^y$ where $c^x_0 = x$, $c^x_t = y$, $c^y_0 = y$, $c^y_t = x$, and $c^x_s = c^y_s$ otherwise.
Note that impatience differs from present bias (Definition 3). Impatience refers to the trade-off between achieving higher satisfaction at earlier rather than later periods; present bias refers to how this trade-off changes when the earlier period occurs in the present rather than in the future.

**Corollary 1.** If axioms 1-5 hold, then $\succ$ exhibits impatience.

Representation (5) with increasing $G$ satisfies further properties in relation to present bias and impatience. First, suppose that at time 0 the agent is indifferent between stream $c$ and $c'$ which involve, at some future $t$, the same trade-off as in Definition 3. Then, he is also indifferent between $s_c$ and $s_{c'}$ in all periods $s$ before $t$. Second, at any $t > 0$, the agent will be impatient and will prefer anticipating at $t$ better future consumption as formally stated in Definition 5. But does he also want to anticipate better consumption at $t$ when considering this possibility before $t$? Because of time inconsistency, the answer can go either way. However, if at $t−1$ the agent does (not) want to anticipate better consumption at $t$, then he also does (not) in all periods before $t$. These properties imply that, under representation (5), it is enough to know how the agent resolves an intertemporal trade-off in the period right before it occurs to know how he does so in all previous periods.

### 6.2 A Bellman-type Equation for Dynamic Choice Problems

To see how to work with representation (5), consider the following consumption-saving problem. For expositional simplicity, we formulate it as a cake-eating problem—it should be clear that the method described here can be generalized to other Markovian decision problems. At 0, the agent must commit to a stream $(c_0, c_1, \ldots) \in \mathbb{R}_+^\mathbb{N}$ subject to the constraint $\sum_{t \geq 0} c_t \leq b$, where $b$ is the cake size. Also, let $C(b)$ be the set of all nonnegative consumption streams satisfying this constraint. Based on representation (5), the optimal utility is given by

$$U^*(b) = \sup_{c_0 \leq b} \left\{ u(c_0) + \alpha W(b - c_0) \right\},$$

where

$$W(b') = \sup_{c' \in C(b')} \sum_{t=0}^{\infty} \alpha^t G(U(tc')).$$

If we can solve for $W$, we can then easily determine the optimal consumption plan. Note that, for any $b \geq 0$, we can express $W(b)$ as

$$W(b) = \sup_{c_0 \leq b} \left\{ \sup_{c' \in C(b-c_0)} \left\{ G \left( u(c_0) + \alpha \sum_{t=0}^{\infty} \alpha^t G(U(tc')) \right) + \alpha \sum_{t=0}^{\infty} \alpha^t G(U(tc')) \right\} \right\}. $$
With increasing $G$, this yields the following Bellman-type equation for $W$:

$$W(b) = \sup_{c_0 \leq b} \{ G(u(c_0) + \alpha W(b - c_0)) + \alpha W(b - c_0) \}$$  \hspace{1cm} (7)

Given $W$, the maximization in (6) determines the optimal $c_0$ and that in (6.2) determines $c_t$ for all $t > 0$.

Equation (6.2) differs from usual Bellman equations mainly because the instantaneous-utility term is inside the function $G$. Indeed, it reduces to a standard equation if $G$ is linear. However, under minor regularity conditions on $G$, equation (6.2) has a well-defined solution $W$. To see this, define the operator $J$ on the set $B(\mathbb{R}_+)$ of bounded real-valued functions of $\mathbb{R}_+$ by

$$J(W)(b) = \sup_{c_0 \leq b} \{ G(u(c_0) + \alpha W(b - c_0)) + \alpha W(b - c_0) \}.$$

Then, if $G$ is bounded and $K$-Lipschitz continuous with $K < (1 - \alpha)/\alpha$, it is easy to show that $J$ is a contraction and therefore has a unique fixed point. So equation (6.2) uniquely defines $W$. It is straightforward to approximate numerically this fixed-point, which is just a univariate function, and the rate of convergence of numerical schemes is given as a function of the Lipschitz constant of $G$.

When the agent cannot commit at 0, time inconsistency leads to an equilibrium problem, in which he chooses $c_t$ at each $t$. Existence and properties of Markovian equilibria in a similar setting—the ‘buffer-stock model,’ which includes stochastic shocks to the state ($b$ here)—have been studied by Ray (1987), Bernheim and Ray (1989), Harris and Laibson (2001), and Quah and Strulovici (2013). Bernheim and Ray study a set of utility functions that includes those in Theorem 2, so their equilibrium analysis applies to the preferences studied here.

### 6.3 Intertemporal Rate of Utility Substitution

Representation (5) has interesting implications on how the agent trades off consumption across periods. By Axiom 2, separability holds between immediate consumption and future well-being, as well as across future well-being. Nonetheless, the trade-off between consumption at 0 and $t > 0$ can depend on well-being—and hence consumption—between 0 and $t$ and after $t$. This is because well-being at $t$ affects well-being at all $s < t$ and depends on well-being after $t$.

To see this, we consider the agent’s discount factor between 0 and $t$. Of course, consumption trade-offs between 0 and $t$ also depend on the curvature of $u$. To bypass this dependence, first note that Theorem 2 suggests that, given $\alpha$ and $G$, the preference is
entirely driven by the instantaneous utility \( u \).

**Corollary 2** (*u*-Representation). Given representation (5) and \( H \)-continuity, there is a continuous nonconstant function \( \hat{U} : \mathbb{R}^T \to \mathbb{R} \) such that, for all \( c \in C \),

\[
U(c) = \hat{U}(u(c_0), u(c_1), \ldots).
\]

Relying on this result, given stream \( c \), define \( u_s = u(c_s) \) and the discount factor as

\[
d(t, c) = \frac{\partial U(c)/\partial u_t}{\partial U(c)/\partial u_0}.
\]

(8)

That is, \( d(t, c) \) is the rate at which the agent substitutes instantaneous utility between 0 and \( t \). Note that, under EDU, \( d(t, c) = \delta^t \). In general, for \( d(t, c) \) to be well defined, the derivatives in (8) must exist. This is always the case when \( G \) is differentiable.\(^{24}\)

**Proposition 4.** Suppose \( G \) in representation (5) is differentiable. Then, \( d(1, c) = \alpha G'(U(1c)) \) and, for \( t > 1 \),

\[
d(t, c) = \alpha^t G'(U(tc)) \left[ 1 + \sum_{\tau=1}^{t-1} G'(U(t-\tau)c) \prod_{s=1}^{\tau-1} (1 + G'(U(t-s)c)) \right],
\]

where \( \prod_{s=1}^{\tau-1} (1 + G'(U(t-s)c)) \equiv 1 \) if \( \tau = 1 \).

This formula has a simple explanation. Suppose \( c_t \) changes so that \( u(c_t) \) rises by a small amount. This has two effects: (1) well-being rises at \( t \), which explains the term \( G'(U(tc)) \); consequently, (2) well-being rises for all \( \tau \) between period 1 and \( t \), which explains the summation. Moreover, the rise in \( U(tc) \) affects \( U(t-\tau c) \) through all well-beings between \( t - \tau \) and \( t \), which explains the product.

By Proposition 4, for general \( G \), the discount factor \( d(t, c) \) depends on consumption at \( t \) as well as on well-being before and after \( t \)—hence it depends on the entire stream \( c \) (see Section 6.5 for details). The only case in which \( d(t, c) \) is independent of \( c \) is when \( G \) is linear. Surprisingly, in this case, the discount factor takes a very well-known form.

**Corollary 3.** Suppose \( G(U) = \gamma U \) with \( 0 < \gamma < \frac{1-\alpha}{\alpha} \). Then, for all \( t > 0 \),

\[
d(t, c) = \beta \delta^t,
\]

where \( \beta = \frac{\gamma}{1+\gamma} \) and \( \delta = (1 + \gamma)\alpha < 1 \).

*Proof.* By Proposition 4, the result is immediate for \( t = 1, 2 \). For \( t > 2 \),

\[
d(t, c) = \alpha^t \gamma \left[ 1 + \gamma \sum_{\tau=1}^{t-1} (1 + \gamma)^{\tau-1} \right] = \alpha^t \gamma (1 + \gamma)^{t-1}.
\]

\(\square\)

\(^{24}\)Note that, when increasing, \( G \) is already differentiable at almost every point in \( U \).
6.4 Quasi-hyperbolic Discounting of Instantaneous Utility

A natural question is which properties of $\succ$ correspond to $G$ being linear and hence to quasi-hyperbolic discounting. As noted, unless $G$ is linear, the trade-off between utility from consumption at 0 and at $t > 0$ depends on well-being between 0 and $t$, and after $t$. This observation suggests Axiom 6.

**Axiom 6 (Trade-off Independence).**

(i) $(c_0, c_1, 2c) \succ (c'_0, c'_1, 2c)$ if and only if $(c_0, c_1, 2c) \succ (c'_0, c'_1, 2c)$;

(ii) $(c_0, c_1, 2c) \succ (c'_0, c'_1, 2c)$ if and only if $(c_0, c'_1, 2c) \succ (c'_0, c'_1, 2c)$.

Intuitively, condition (i) says that the trade-off between consumption at 0 and at 1 is independent of the continuation stream $2c$ and hence of future well-being. Condition (ii) says that the trade-off between consumption at 0 and after 1 is independent of consumption at 1 and hence of well-being at 1.

**Theorem 3 (‘Vividness’ Well-being Representation).** Axiom 1-6 hold if and only if the function $U$ may be chosen so that

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

(9)

where $\alpha \in (0, 1)$, $\gamma \in (0, \frac{1-\alpha}{\alpha})$, and $u : X \to \mathbb{R}$ is a continuous nonconstant function.

**Corollary 4 (Quasi-Hyperbolic Discounting).** Axiom 1-6 hold if and only if there are $\beta, \delta \in (0, 1)$ and a continuous nonconstant function $u$ such that

$$U(c) = u(c_0) + \beta \sum_{t=1}^{\infty} \delta^t u(c_t).$$

**Proof.** By Theorem 3, for all $t$, $U(c)$ is a strictly increasing, linear function of $U(tc)$, which is in turn a strictly increasing, linear function of $u(c_t)$. Hence, there is a function $\kappa(t) : T \setminus \{0\} \to \mathbb{R}_{++}$ such that

$$U(c) = u(c_0) + \sum_{t=1}^{\infty} \kappa(t) u(c_t).$$

Clearly, for all $t > 0$, $\kappa(t) = d(t, c)$ defined in (8). Corollary 3 implies the result. □

This result allows us to understand $\beta-\delta$ discounting of instantaneous utility in terms of simple properties of fully forward-looking preferences. First, a $\beta-\delta$ agent treats immediate consumption and future well-being as different entities. This explains the stark difference between discounting of future instantaneous utility from today’s perspective and from any future period’s perspective, which seems ad hoc and otherwise hard to justify. Second,
a $\beta$-$\delta$ agent cares directly about his well-being in all future periods, in a stationary way. This explains present bias, namely $\beta < 1$. Finally, he treats in a separable way immediate consumption and future well-being as well as well-being across future periods, and trades off consumption between any two periods in a way that does not depend on intermediate and future well-being. This explains additive time separability in instantaneous utility.

This explanation of quasi-hyperbolic discounting differs from the usual one based on myopia, which views the agent as caring disproportionately about the present against any future period—if anything, fully forward-looking behavior captures the opposite. Moreover, Corollary 4 tightly links the degree of present bias, $\beta$, with the agent’s ex-ante marginal utility from future well-being, $\gamma$. We can also interpret $\gamma$ as measuring the degree to which the agent finds future outcomes ‘imaginable’ or ‘vivid.’ This vividness interpretation relates to Böhm-Bawerk’s (1890) and Fisher’s (1930) idea that an agent’s current utility depends on immediate consumption as well as on his ability to imagine or foresee his future ‘wants.’

Corollary 4 has other implications. First, it can explain why present bias may weaken as the period length shortens. If each period represents a shorter time horizon—a week rather than a month—current instantaneous utility $u(c_0)$ should play a smaller role in determining current well-being; that is, $\gamma$ should be larger in (9). Consequently, $\beta$ should get closer to one. The second implication is about long-run discount factors and how increasing vividness of future well-being can mitigate present bias.\footnote{Vividness of the future well-being implied by today’s decisions may be influenced with specific ad campaigns. For example, consider the dramatic pictures and reminders printed on cigarette packs. It is hard to believe that such packaging is just meant to inform unaware customers of the consequences of smoking.}

**Corollary 5.** For any well-being discount factor $\alpha$, increasing vividness $\gamma$ mitigates the agent’s present bias and makes him discount less instantaneous utility in the long run: both $\beta$ and $\delta$ increase. For any degree of present bias $\beta$, the agent discounts less instantaneous utility than future well-being in the long run: $\delta > \alpha$.

For example, using Laibson et al.’s (2007) 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.

### 6.5 Beyond Quasi-hyperbolic Discounting

Representation (9) is appealing, for it corresponds to a widely used and tractable model which captures well-documented behavioral phenomena like present bias. Quasi-hyperbolic
discounting, however, cannot capture other phenomena that appear as anomalies through the lenses of EDU. Instead, representation (5) can capture and, most importantly, can offer a general explanation for some of these anomalies.

First of all, some agents discount future consumption at a rate that depends on the level of consumption itself (for evidence, see Frederick et al. (2002)). By Proposition 4, this phenomenon would occur because how much an agent discounts instantaneous utility from $c_t$ depends on how much his ex-ante utility from well-being at $t$ (i.e., $G(U(c_t))$) responds to changes in $c_t$, which may depend on the level of $c_t$. More generally, $d(t, c)$ can depend on the entire stream $c$. This may capture Fisher’s (1930) idea that “[an agent’s] degree of impatience for, say, $\$100$ worth of this year’s income over $\$100$ worth of next year income depends upon the entire character of his […] income stream pictured as beginning today and extending into the indefinite future.” (Ch. 4, §5)

How $d(t, c)$ varies with $c$ ultimately depends on the properties of $G$. If $G'$ is decreasing (increasing), then the agent discounts more a stream that yields higher (lower) instantaneous utility in all periods. To state this result formally, for any $c, c' \in C$, let $u(c_t) \geq u(c'_t)$ if and only if $u(c) \geq u(c')$ for all $t \geq 0$.

**Corollary 6.** Let $d(t, c)$ be as in Proposition 4. For any $t > 0$, $c \geq_u c'$ implies $d(t, c) \leq d(t, c')$ if and only if $G'$ is decreasing. Conversely, $c \geq_u c'$ implies $d(t, c) \geq d(t, c')$ if and only if $G'$ is increasing.

This result shows a tight link between discounting and a property of $G$, which in principle can be empirically tested.

According to some evidence, intertemporal preferences can also exhibit consumption interdependences across dates, in contrast with both EDU and the $\beta$-$\delta$ model (again, see Frederick et al. (2002)). Clearly, a general non-separable representation like that in Assumption 2 allows for such interdependences, but offers no insight on their origin. By contrast, representation (5) identifies a specific source of interdependences. Although the agent treats each period separately with regard to well-being, how he evaluates consumption at a future $t$ can depend on well-being, and hence consumption, before and after $t$. This can explain forward- and backward-looking interdependences that arise only in the future.\footnote{Such interdependences differ from other backward-looking interdependences, like habit formation, which this paper cannot capture.}

Such interdependences can create a preference for enjoying similar consumption events far apart over time, in a way consistent with some evidence. For example, Frederick
et al. (2002) discuss the following experiment (p. 364). There are five periods and two opportunities, A and B, to dine at a fancy restaurant. Subjects are asked to rank consumption streams in two scenarios. In scenario 1, one stream features A at time 0 and the other features A at time 2; in all other periods, both streams involve dining at home. Scenario 2 coincides with scenario 1, except that both streams feature B at time 4. Presumably, dining at a fancy restaurant yields higher instantaneous utility than dining at home. According to the experiment, in scenario 1, most of the subjects prefers having A at 2. But in scenario 2 a significant fraction of them prefers having A at 0. This change is consistent with a concave $G$—perhaps the most natural case, since it simply means that the marginal value at 0 of well-being at any $t > 0$ is decreasing. In scenario 1, an agent may prefer to delay A, for the improvement in well-being both at 2 and at 1 may offset discounting. But, when well-being at 4 is higher because of B, well-being at 2 is also higher. Consequently, the benefit of improving time-2 well-being by delaying A is lower than when B is absent; hence the agent prefers enjoying A right away.

Finally, representation (5) can explain other empirical findings summarized by Frederick et al. (2002). The sign effect, for example, means that gains are discounted more than losses. When $G$ is concave, increasing consumption at some time $t$ reduces the discount factor between 0 and $t$. Hence, from EDU’s viewpoint, it appears as if the discount rate increases. Intuitively, the discount factor captures an indifference point in the trade-off between instantaneous utility at 0 and at $t$. If such utility at $t$ is higher, this indifference occurs at a lower point, for the time-0 marginal value of well-being at $t$ is lower—in a sense, improving future well-being matters less.

7 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) 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 $\delta$ (which is independent of $\beta$).” (p. 1208)

27In infinite-horizon settings, by Corollary 1, the agent never wants to delay delightful events. This can happen, however, in a finite-horizon version of representation (5) (see Appendix A).
Another, perhaps more fundamental, issue is whether the existence of agents with time-inconsistent preferences justifies some form of paternalistic interventions. An immediate consequence of the present paper is to weaken the case for such interventions. If time inconsistency were the result of some bounded rationality, one might be tempted to argue that time-inconsistent agents can benefit from paternalistic interventions. This argument, however, is not valid if time inconsistency is the result of fully forward-looking behavior. In this case, why should a planner use welfare criteria other than the agent’s ex-ante preference? If at 0 he is already fully taking into account his future preferences over continuation streams, then the planner may simply adopt a ‘libertarian’ stance and use $W^L(c) = U(c)$ to measure welfare.

From this paper’s viewpoint, a libertarian criterion seems even more appropriate for time-inconsistent agents than for time-consistent ones. Indeed, under EDU, welfare is usually and uncontroversially measured using the agent’s current preference. Such preference, however, takes into account his preferences only in the next period, but not in subsequent ones. By contrast, the $\beta$-$\delta$ model implies that the current preference takes into account, albeit in a simple way, preferences in all future periods—of course, this also holds for more general, fully forward-looking preferences. In this view, the ex-ante preference implied by the $\beta$-$\delta$ model may seem a more reasonable welfare criterion than that implied by EDU. For example, consider two policies, $A$ and $B$, inducing streams $c^A$ and $c^B$ such that $1c^A \sim^1 1c^B$ but $t_c^A \succ^t t_c^B$ for all $t > 1$. The criterion based on EDU implies that at 0 $A$ is as desirable as $B$. Instead, the criterion based on the $\beta$-$\delta$ model implies that at 0 $A$ is strictly more desirable. Thus, the second criterion favors what we may call long-run sustainability and may therefore seem more appealing.

Using the well-being representation in Theorem 3, we can also derive the welfare criterion that the literature has so far assumed for the $\beta$-$\delta$ model. Recall that ex-ante choices of a $\beta$-$\delta$ agent reveal the corresponding parameters $\alpha$ and $\gamma$. We may interpret the factor $\alpha_t^t$ as capturing the agent’s assessment of how likely it is that a policy under scrutiny will continue to matter at $t$. This assessment may combine subjective aspects as well as objective information, which the agent may know better than the planner. Therefore, the planner may use the weights $\alpha_t^t$ to aggregate the agent’s well-being across periods. Surprisingly, doing so is equivalent to aggregating his instantaneous utilities using the familiar weights $\delta^t$. Thus, a natural, focal specification of welfare weights for a time-inconsistent agent delivers a time-consistent planner.

**Proposition 5.** Suppose $U(c)$ can be represented as in Theorem 3 and Corollary 4, with corresponding parameters $(u, \alpha, \gamma)$ and $(u, \beta, \delta)$. Let $w : T \rightarrow \mathbb{R}_+$ and define
$W(c) = \sum_{t=0}^{\infty} w(t)U(t)$. Then,

$$W(c) = \sum_{t=0}^{\infty} \delta^t u(c_t)$$

if and only if $w(t) = \alpha^t$.

One direction of this result follows from an intuitive argument. Let $W(c) = \sum_{t=0}^{\infty} \alpha^t U(t)$. Using $\alpha - \gamma$ representation of $U(c)$, we have

$$W(c) = u(c_0) + (1 + \gamma)\alpha \sum_{t \geq 1} \alpha^{t-1} U(t) = u(c_0) + (1 + \gamma)\alpha W(1c).$$

This shows why the planner is time consistent and $W(c)$ corresponds to the sum of instantaneous utilities exponentially discounted with factor $\hat{\delta} = (1 + \gamma)/\alpha$. Intuitively, the planner is time consistent because she values well-being from time 1 onward in the same way as does the agent, but for her this value coincides with her ‘continuation utility’ $W(1c)$. Of course, we know from Corollaries 3 and 4 that $\hat{\delta} = \delta$ in the $\beta-\delta$ version of $U(c)$. However, to see why this has to be the case for any $\hat{\delta}$, note that the $\alpha-\gamma$ representation satisfies

$$U(c) = u(c_0) + \gamma \alpha W(1c) = u(c_0) + \frac{\gamma \alpha}{\delta} \sum_{t>0} \hat{\delta}^t u(c_t).$$

It follows that $U(c)$ must have a quasi-hyperbolic representation in terms of instantaneous utilities, where the long-run discount factor coincides with the planner’s factor.

Since time consistency may reveal that an agent is myopically forward-looking, a natural question is then whether the planner should weigh more his well-being beyond the immediate future than does the agent himself and, if so, by how much. In the case of EDU, i.e., $U(t) = \sum_{t=0}^{\infty} \delta^t u(c_t)$, one might rely on $\delta$ to aggregate well-being across periods using the criterion $\hat{W}(c) = \sum_{t=0}^{\infty} \delta^t U(t)$ (see, e.g., Ray (2014)). Doing so, however, makes the planner time inconsistent. Indeed, one can show that $\hat{W}(c) = u(c_0) + \sum_{t=1}^{\infty} \delta^t (1 + t) u(c_t)$.\footnote{At first glance, EDU appears as the limit of the $\beta-\delta$ model as $\beta \to 1$. However, by Theorem 3 and Corollary 3, in the limit the corresponding $\alpha$ and $\gamma$ must take extreme, implausible values. Therefore, it seems more natural to think of time-consistent and time-inconsistent agents as having radically different preferences.}

To summarize, for the $\beta-\delta$ model discussed by Rubinstein, this paper shows that the arguably most natural, welfare criteria are precisely those which have been used in practice. The paper also sheds light on their respective implications on how the planner weighs the agent’s well-being across periods.

\footnote{See Appendix B for the proof of the other direction.}
8 General Discounting: Well-being vs. Instantaneous Utility

Discounting is a standard feature in dynamic economic models. This section investigates the general relationship between how the agent discounts future instantaneous utility—e.g., in the case of hyperbolic discounting—and how he cares about future well-being.

In general, a discount function captures the weight that, at $t$, the agent assigns to his instantaneous utility at $s > t$. For $s > t \geq 0$, let this weight be $d(t, s) \in (0, 1)$. Also, suppose that his preference can be represented by

$$U_t(t, c) = u(c_t) + \sum_{s > t} d(t, s)u(c_s)$$

for some function $u : X \to \mathbb{R}$.\(^{30}\) The questions here are whether we can always find a weight function $q$ such that

$$U_t(t, c) = u(c_t) + \sum_{s > t} q(t, s)U_s(s, c),$$

and how $q$ relates to $d$. To state the answer, for $s > t$, let $\mathcal{T}(t, s)$ be the set of all increasing vectors of dates starting at $t$ and ending at $s$ and $\hat{\mathcal{T}}(t, s)$ the set of all such vectors with at least one intermediate date:

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

and for $s > t + 1$

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

**Proposition 6.**

(i) If $U_t(t, c)$ satisfies (10) for some discount function $d$, then it also satisfies (11) for some weight function $q$. In particular, for all $t \geq 0$ and $s > t + 1$, $q(t, t + 1) = d(t, t + 1)$ and

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

(ii) If $U_t(t, c)$ satisfies (11) for some non-negative weight function $q$, then it also satisfies (10) for some discount function $d$. Moreover, for all $t \geq 0$ and $s > t$,

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

Proposition 6 immediately implies that, independently of how the agent weighs future well-being, he weighs instantaneous utility at least two periods in the future strictly more.

---

\(^{30}\)This section allows for the possibility that preferences are not time invariant (Assumption 1).
than his well-being at that period.

**Corollary 7.** Suppose that \( d \) and \( q \) are the instantaneous-utility discount function and the well-being weight function corresponding to the same \( \succ \). Then, \( d(t, s) > q(t, s) \) whenever \( s - t \geq 2 \).

Using Proposition 6, we can also derive which properties of the general discount function \( d \) are equivalent to myopically forward-looking behavior, namely \( q(t, s) = 0 \) for \( s > t + 1 \). Note that in all cases \( d(t, t + 1) = q(t, t + 1) \).

**Corollary 8.** For all \( t \geq 0 \), \( U_t(c) = u(c_t) + d(t, t+1)U_{t+1}(c_{t+1}) \) if and only if, for \( s > t \),

\[
d(t, s) = \prod_{j=0}^{s-t-1} d(t+j, t+j+1).
\]

Corollary 8 implies that we can interpret any form of discounting that depends only on the time distance \( s - t \) and is not exponential (like hyperbolic discounting) as revealing that, when comparing consumption streams, the agent cares directly about well-being beyond the immediate future—i.e., he is not myopically forward-looking. This is because, if \( d \) depends only on \( s - t \), then of course we must have exponential discounting: Letting \( \delta = d(1) \), by (16) we get \( d(s, t) = \delta^{s-t} \).

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

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

with \( k, p > 0 \) (see Loewenstein and Prelec (1992)). Unfortunately, it is hard to derive in closed form the well-being weight function \( q^h \) corresponding to \( d^h \).

However, using Proposition 6, we can simulate \( q^h \) for different values of \( k \) and \( p \). This gives us a qualitative picture of how \( q^h \) relates to \( d^h \).

Figure 1 represents \( d^h \), \( q^h \), and the exponential discount function for \( \delta = 0.9 \) (curve ED). 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, given \( p \), \( d^h(\tau) \) converges to \( \delta^\tau \) for every \( \tau \) as \( k \to 0 \)—this pattern is clear going from panel (a) to (c). All three panels illustrate the point of Corollary 7: Starting at time 2, \( q^h \) is strictly below \( d^h \). Moreover,

---

\(^{31}\)Using a recursive formulation, Saez-Marti and Weibull (2005) make a similar but weaker observation, namely that \( d(t, s) \geq q(t, s) \) for all \( t \) and \( s \).

\(^{32}\)Proposition 2 in Saez-Marti and Weibull (2005) implies that, if \( d^h \) satisfies (17), then \( q^h \geq 0 \).
Figure 1: Hyperbolic-Discounting Simulation

(a) \( k = 30 \)

(b) \( k = 0.3 \)

(c) \( k = 0.0001 \)
they illustrate that, when a hyperbolically-discounting agent becomes more similar to an exponentially-discounting agent, he discounts more well-being from period 2 onwards: For these periods, $q^h$ moves closer to zero as $d^h$ moves closer to ED. This is, of course, an illustration of Propositions 1 and Corollary 8: the agent can be time consistent if and only if he does not care about well-being beyond time 1, i.e., $q^h(\tau) = 0$ for $\tau > 1$.

9 Relation to Other Stationarity Notions

This section briefly compares well-being stationarity (Axiom 4) with Koopmans’s (1960, 1964) stationarity and Olea and Strzalecki’s (2013) quasi-stationarity.

In his general analysis of impatience, Koopmans considers a single preference, $\succeq^0$, and calls it stationary if it satisfies the following property: If $1c \sim^0 1c'$, then $(c_0, 1c) \sim^0 (c_0, 1c')$; and if $1c \succ^0 1c'$, then $(c_0, 1c) \succ^0 (c_0, 1c')$. This property looks similar to time consistency (Definition 2), but is conceptually very different. 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." (Koopmans [1964], p. 85, emphasis in the original)

The point can be illustrated with two simple examples.

**Example 1** (Stationarity $\nRightarrow$ Time Consistency): For $t \geq 0$, $\succ^t$ is represented by

$$U_t(tc) = u(c_t) + \sum_{s>t}^{\infty} \delta_s^{t-s}u(c_s) = u(c_t) + \delta_tU_{t+1}(c),$$

where $u : X \to \mathbb{R}$ is continuous and strictly increasing. Moreover, $\delta_t \in (0, 1)$ and $\delta_t > \delta_{t+1}$ for $t \geq 0$. Then, each $\succ^t$ satisfies stationarity, but $\{\succ^t\}_{t \geq 0}$ violates time consistency.

**Example 2** (Time Consistency $\nRightarrow$ Stationarity): For $t \geq 0$, $\succ^t$ is represented by

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

where $\delta \in (0, 1)$, $\phi_0 \in (0, 1)$, and $\phi_t = 1$ for $t > 0$; $u(\cdot)$ is as before. Then $\{\succ^t\}_{t \geq 0}$ satisfies time consistency, but $\succ^0$ violates stationarity.

---

\textsuperscript{33} Of course, if Assumption 1 (Time Invariance) holds, time consistency and stationarity are mathematically equivalent.
Using another axiom—i.e., separability of $>0$ in $c_0$ and $1c$—Koopmans obtains the recursive representation $U(c) = \hat{V}(u(c_0), U(1c))$, where $\hat{V}$ is strictly increasing in each argument. Note that separability implies that $U(c)$ depends on $1c$ only through a ‘perspective utility’ $U_1(1c)$, which equals $U(1c)$ by stationarity. With further axioms, Koopmans shows that there exists a specific form of $\hat{V}$ which corresponds to EDU. The previous remarks and Proposition 1 remind us that EDU features time consistency not because it discounts future instantaneous utility at a constant rate—a cardinal property of a specific $\hat{V}$—but because today’s well-being depends only on tomorrow’s well-being—an ordinal property of $\hat{V}$.34

If we view the agent as evaluating streams based only on the instantaneous utility generated in each period—so that consumption at 0 and at $t > 0$ are conceptually equivalent—then Koopmans’s stationarity seems reasonable. If we delay two streams from 0 to 1, replacing consumption at 0 with the same level in both cases, why should the agent, at 0, rank the new streams differently from the old ones? Note that, with representation $\hat{V}$, stationarity holds between any two periods.

Olea and Strzalecki’s paper allows for violations of stationarity, but only between period 0 and 1—the resulting property is called quasi-stationarity. Since their paper continues to view the agent as evaluating streams based only on instantaneous utilities, quasi-stationarity seems more difficult to justify. If the agent 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 with well-being stationarity (Axiom 4), for in the present model tomorrow’s well-being is equivalent to well-being thereafter, but is essentially different from today’s consumption.

These remarks clarify how the present paper departs from other intertemporal utility models. It does not relax stationarity directly—for example, as Hayashi (2003) and Olea and Strzalecki (2013) do. Instead, it takes a conceptually different view of what determines preferences over consumption streams. The basic 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 Koopmans’s stationarity.

34A property of a representation is ordinal if it is invariant to strictly increasing monotonic transformations.
10 Conclusion

This paper has developed a theory of forward-looking preferences in which agents directly incorporate their preferences at future dates to determine their current preference over consumption streams. The resulting preferences have a tractable representation and makes precise predictions regarding agents’ behavior. Present bias and consumption dependence, which were relegated to the status of anomalies under the exponentially-discounted-utility paradigm, are natural consequences of fully forward-looking preferences, and the theory sheds a new light on those concepts. For example, present bias may have nothing to do with an “excessive” taste for the present. As a byproduct, the theory provides an axiomatization of the renowned quasi-hyperbolic discounting model, which is explained by fully forward-looking agents whose future preferences enter “linearly” in their current preference.

Despite depending recursively on preferences at future periods, any fully forward-looking preference exhibits impatience. Koopmans’s (1960) axiomatization of EDU implies a strictly positive discount rate and, hence, well-defined preferences over an infinite horizon. Similarly, this paper’s axiomatization implies that future preferences are discounted positively and, under a stationarity axiom, that variations in future well-being (under an intuitive interpretation of the model) have limited effects on current well-being. This implies, for instance, that fully forward-looking agents cannot become infinitely happy by anticipating future pleasurable events. This feature, among others, distinguish the theory from previous ad-hoc models of anticipations, in which parameter restrictions ensuring that the model was well-defined were left unjustified.

The theory allows us to conduct rigorous welfare analysis with time-inconsistent agents of the type studied here. Paternalistic interventions seem difficult to justify, since fully forward-looking agents take into account their well-being in all future periods. For instance, these agents are never victim of money-pump schemes, for they fully anticipate how such schemes would affect their future well-being. These agents may, however, demand commitment and perfectly predict that, in the future, they will want to walk away from it. Any justification of policies helping time-inconsistent agents honor their commitments ultimately hinges on how we measure welfare. Welfare criteria that provide such a justification have been usually assumed in the literature. The paper shows that we can derive such criteria from representations of fully forward-looking preferences.
Time-inconsistent Preferences vs. Time-inconsistent Behavior

Time inconsistency is sometimes regarded as unimportant for economic analysis.\footnote{See, e.g., Mulligan (1996).} One argument in support of this view is that a time-inconsistent agent would be subject to money-pump schemes which would quickly lead him to complete immiseration. Anticipating this, the agent would never trade in markets. Consequently, there would be no reason to allow for time-inconsistent agents in economic models.

This view, like ours, emphasizes the importance to distinguish time inconsistency of preferences with time inconsistency of actual choices. Needless to say, these two concepts should be kept separate. As already suggested by Strotz (1955), there seem to be at least two reasons why time-inconsistent preferences need not result in time-inconsistent choices. Fully aware of his inconsistency, an agent may anticipate his future inability to carry out certain decisions. Hence, on the one hand, he may commit to them in advance (Strotz’s ‘strategic precommitment’). On the other hand, he may consider only courses of action that he can implement over time (Strotz’s ‘consistent planning’). So, depending on the economic environment, it seems perfectly conceivable to observe agents with time-inconsistent preferences make time-consistent choices. This does not imply, of course, that we could never distinguish consistent from inconsistent agents—only the latter, for instance, value commitment ex ante. Moreover, the sort of intra-personal equilibrium that a time inconsistent agent may engage in with his future selves may have different properties from the consumption path chosen for a time consistent agent.\footnote{See Harris and Laibson (2001).} For example, there may exist multiple equilibria, and this multiplicity may lead to choices that vary discontinuously in parameters of the model. The analysis of intertemporal choice is obviously completely different if the agent has time-inconsistent preferences, and one cannot simply focus on time-consistent agents on the ground that this analysis may lead to time-consistent behavior.

These considerations and the paper’s analysis may shake our confidence in the standard EDU model, often based on its property of offering a time-consistent choice rule. As noted, to meaningfully elicit preferences over consumption streams from how the agent chooses among them, we must assume that he is committing to them. But if there is commitment, it is unclear why time consistency is a ‘virtue’ of EDU. Without commitment, time consistency of the EDU model may become appealing to the agent. But without commitment, the agent is obviously unable to commit to obeying that model. Instead, one has to consider intrapersonal equilibria, in which the agent is making plans under the concern that he will revise them in the future. This possibility, however, adulterates
the exercise of eliciting the agent’s underlying preference and says nothing about whether EDU is a good model of such preference.

Finally, this paper contributes to the discussion on which behaviors can be classified as resulting from ‘bounded rationality.’ Indeed, it shows that a form of behavior—time inconsistency—that may seem to fall into this category can be explained by a model of fully rational decision making. Not all preferences are like that. It would also be natural to extend the theory to allow for backward-looking preferences, allowing agents to include past consumption in their current preference. While such extension is of clear interest, it raises quite distinct issues which we leave for future research.
11 Appendix A: Anticipating Dreadful Events and Delaying Delightful Ones.

This section illustrates that, in a finite-horizon setting, a well-being representation similar to that in Theorem 2 can accommodate a desire to anticipate dreadful events and to delay delightful ones. Let $T = \{0, 1, \ldots, \bar{t}\}$ with $\bar{t} < \infty$, $\overline{C} = X^{\bar{t}+1}$, and $\alpha \in (0, 1)$, $u$, and $G$ as in Theorem 2. We can recursively define a well-being representation as follows. For any $c \in \overline{C}$, let $U_t(c) = u(c)$ and, for $t < \bar{t}$, let

$$U_t(c) = u(c) + \sum_{\tau=1}^{\bar{t}-t} \alpha^\tau G(U_{t+\tau}(c_{t+\tau})).$$  \hfill (18)

Since $U_t$ is constructed by backward induction for each $t$, it is well defined even if $\alpha$ and $G$ do not satisfy the restrictions in Proposition 3. Hereafter, suppose $G$ is strictly increasing. It is also easy to see—adapting Proposition 4 and Corollary 3—that, if $G(U) = \gamma U$ with $\gamma > 0$, then

$$U_t(c) = u(c) + \sum_{\tau=1}^{\bar{t}-t} \beta^\tau u(c_{t+\tau}),$$

where $\beta = \frac{\gamma}{1+\gamma}$ and $\delta = \alpha(1+\gamma)$. Note that, for any $\gamma > 0$, $\beta < 1$; yet, it is possible that $\delta > 1$.

Consider now the effect of delaying an event that involves changing consumption from $y$ to $x$. Let $\overline{\pi} = u(x) - u(y)$—e.g., for a dreadful event $\overline{\pi} > 0$. If the agent delays this event from 0 to $t$, his instantaneous utility changes by $\overline{\pi}$ at 0 and by $-\overline{\pi}$ at $t$. Consequently, his well-being changes at $t$ and, indirectly, at all $\tau$ between 1 and $t$. Formally, we can compute the overall effect of an event delay as follows. For $\tau = 1, \ldots, t$, define $\Delta G_{t-\tau}$ recursively as

$$\Delta G_{t-\tau} = G\left(U_{t-\tau}(c_{t-\tau}) + \sum_{k=1}^{t-\tau} \alpha^k \Delta G_{t-s+k}\right) - G(U_{t-\tau}(c_{t-\tau})).$$

The overall effect of the delay until $t$ on well-being at 0 is

$$\theta_t(\overline{\pi}) = \overline{\pi} + \Gamma_t(\overline{\pi}),$$

where

$$\Gamma_t \equiv \sum_{s=0}^{t-1} \alpha^{t-s} \Delta G_{t-s}.$$

So, the agent wants to delay an event from 0 to $t$ if and only if $\theta_t(\overline{\pi}) > 0$. In this case, he also wants to choose $t$ so as to maximize $\theta_t(\overline{\pi})$.

Even though the agent discounts future utility from well-being ($\alpha < 1$), he may prefer
Figure 2: Anticipating and Delaying Consumption Events

to delay delightful events or to anticipate dreadful ones. Note that, in general, $\Gamma_t$ depends
on the stream $\{U_\tau\}_{\tau=1}^t$ of well-being that the agent expects up to period $t$—unless $G$ is
linear or $U_\tau$ is constant and $\bar{\pi}$ is infinitesimal. A complete investigation of this dependence
is beyond the scope of this section, which therefore focuses on a simulation. Suppose
that $G(U) = -e^{-gU}$ with $g > 0$ and $U_\tau = U_0(1 + r)^\tau$ for some $r \in [0, 1]$ and $U_0 > 0$.
Figure 2 reports $\theta_t(\bar{\pi})$ for different specifications of the parameters and $\bar{\pi}$.

Panel (a) considers a delightful event ($\bar{\pi} < 0$) and different well-being streams $\{U_\tau\}_{\tau=1}^t$—
note that the higher $U_0$, the higher the entire stream. When the agent faces a relatively
low well-being stream ($U_0 = 0.1$), he prefers to delay the event to some period in the near future, beyond period 1. This is because, for low $U_\tau$, $G$ responds enough to improvements
in well-being, so that the total effect of the benefit $-\bar{u}$ at some period in the near future can prevail over discounting. However, when facing higher streams, the agent prefers to delay the event to a closer period ($U_0 = 0.105$) or not to delay it at all ($U_0 = 0.115$). This is because, for higher $U_\tau$, concavity makes $G$ less responsive to the benefit $-\bar{u}$. Finally, in all cases, eventually discounting dominates, so that the agent does not want to delay a delightful event indefinitely.\[37\]

Panel (b) considers a dreadful event ($\bar{u} > 0$) and the same well-being streams as in panel (a). If possible, the agent would always want to delay the event indefinitely.\[38\] This is intuitive: Eventually discounting dominates, killing any effect of future well-being losses on current well-being. However, in the short run, the agent prefers to delay the event to period 1, than to any period right after 1. This non-monotonic preference arises because, in the short run, $G$’s sensitivity to well-being losses due to $-\bar{u}$ dominates on discounting. Therefore, anticipating $-\bar{u}$ from, say, period 3 to period 1 allows the agent to avoid its negative effects in periods 1 and 2.

Panel (c) focuses on the effect of discounting on the agent’s desire to delay delightful events—the case of dreadful events is similar. The agent wants to delay as long as discounting is not too strong. This is intuitive: He is willing to delay as long as his current well-being can benefit from the effect, on intermediate well-being, of increasing well-being in a future period by $-\bar{u} > 0$. As $\alpha$ decreases, the agent discounts this cumulative benefit more and consequently is less willing to forgo enjoying the event right away.

Finally, panel (d) focuses on the effect of $G$’s sensitivity on the agent’s desire to delay delightful events—again, the case of dreadful events is similar. The agent wants to delay as long as $G$ is sensitive enough to the benefit $-\bar{u} > 0$ on future well-being. Note that $g$ measures $G$’s elasticity normalized by $U$. Hence, as $g$ increases, the benefits of increasing well-being in the near future by $-\bar{u}$ are more likely to dominate discounting, thus causing delay.

\[37\]The agent would want to delay indefinitely if $G$ were linear and $\delta = \alpha(1 + \gamma) > 1$, which seems unreasonable.
\[38\]The agent would never want to delay indefinitely if $G$ were linear and $\delta = \alpha(1 + \gamma) > 1$, which again seems unreasonable.
12 Appendix B: Omitted Proofs

12.1 Existence of a Utility Representation

The following axioms are standard.

**Axiom 7 (Weak Order), \(\succ\) is complete and transitive.**

For all \(c \in C\), the sets \(\{c' \in C : c' \prec c\}\) and \(\{c' \in C : c' \succ c\}\) are open.

For all \(c \in C\), there are constant sequences \(\tau\) and \(\tau'\) such that \(\tau \preceq c \preceq \tau'\).

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

**Theorem 4.** Under Axioms 7-7, there is a continuous function \(U : C \to \mathbb{R}\) such that \(c \succ c'\) if and only if \(U(c) > U(c')\).

**Proof.** The proof follows and generalizes that of Diamond (1965), and is based on the following lemma by Debreu (1954).

**Lemma 1.** Let \(C\) be a completely ordered set and \(Z = (z_0, z_1, ...)\) be a countable subset of \(C\). If for every \(c, c' \in C\) such that \(c \prec c'\), there is \(z \in Z\) such that \(c \preceq z \preceq c'\), then there exists on \(C\) a real, order-preserving function, continuous in any natural topology.\(^{39}\)

**Lemma 2.** For any \(c \in C\), there is a constant stream \(c^*\) such that \(c \sim c^*\).

**Proof.** Let \(D\) be the set of constant streams and, for any fixed \(c \in C\), let \(A = \{d \in D : d \preceq c\}\) and \(B = \{d \in D : d \succeq c\}\). By Axiom 7, \(A \cup B = D\); by Axiom 7, \(A\) and \(B\) are closed; by Axiom 7, \(A\) and \(B\) are nonempty. Moreover, \(D\) is connected. Indeed, for any continuous function \(\phi : D \to \{0, 1\}\), the function \(\tilde{\phi} : X \to \{0, 1\}\) defined by \(\tilde{\phi}(x) = \phi(x, x, ...)\) is also continuous. Connectedness of \(X\) implies that \(\tilde{\phi}\) is constant and, hence, that \(\phi\) is constant, showing connectedness of \(D\). This implies that \(A \cap B \neq \emptyset\).

To conclude the proof of Theorem 4, let \(Z_0\) denote a countable dense subset of \(X\), which exists since \(X\) is separable, and let \(Z\) denote the subset \(C\) 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.

The next axiom ensures that \(U\) is nonconstant in the first and some other argument.

\(^{39}\)A natural topology is one under which Axiom 7 holds for that topology.
**Axiom 8 (Non triviality).** There are \( x, x', \hat{x} \in X \) and \( c,c',\hat{c} \in C \), such that
\[
(x,\hat{c}) \succ (x',\hat{c}) \text{ and } (\hat{x},c) \succ (\hat{x},c').
\]

### 12.2 Proof of Proposition 1

Suppose that \( V(c_0,U(1c),U(2c),\ldots) = V(c_0,U(1c)) \) for all \( c \in C \) and \( V \) is strictly increasing in \( U(1c) \). By Assumption 1, if \( 1c \sim^1 1c' \), then \( U(1c) = U(1c') \) and, since \( V \) is a function, \( V(c_0,U(1c)) = V(c_0,U(1c')) \); hence \( (c_0,1c) \sim^0 (c_0,1c') \). If \( 1c \succ^1 1c' \), then \( U(1c) > U(1c') \) and, since \( V \) is strictly increasing in its second argument, \( V(c_0,U(1c)) > V(c_0,U(1c')) \); hence \( (c_0,1c) \succ^0 (c_0,1c') \).

Suppose \( 1c \sim^1 1c' \) implies \( (c_0,1c) \sim^0 (c_0,1c') \). Then, for any \( (U(1c),U(2c),\ldots) \) and \( (U(1c'),U(2c'),\ldots) \) such that \( U(1c) = U(1c') \),
\[
V(c_0,U(1c),U(2c),\ldots) = V(c_0,U(1c'),U(2c'),\ldots).
\]
So \( V \) can depend only on its first two arguments. Suppose \( 1c \succ^1 1c' \) implies \( (c_0,1c) \succ^0 (c_0,1c') \). Then, \( U(1c) > U(1c') \). Moreover, it must be that \( V(c_0,U(1c)) > V(c_0,U(1c')) \); that is, \( V \) must be strictly increasing in its second argument.

### 12.3 Proof of Proposition 2

Let \( U(c) = V(c_0,U(1c),U(2c),\ldots) \) where \( V \) is strictly increasing in \( U(tc) \) for all \( t > 0 \). By definition, \( (x,c) \succ (y,c) \) means that \( U(x,c) > U(y,c) \). Hence, for all \( 0 \leq s \leq t \),
\[
U(sz_t,x,c) > U(sz_t,y,c),
\]
where, for \( s < t \), \( sz_t = (z_s,\ldots,z_t) \) and \( tz_t = z_t \). This follows by induction. For \( s = t \),
\[
U(tz_t,x,c) = V(tz_t,U(x,c),U(c),\ldots) > V(tz_t,U(y,c),U(c),\ldots) = U(tz_t,y,c).
\]
Now suppose that the claim holds for \( r + 1 \leq s \leq t \), with \( 0 \leq r < t \). Then
\[
U(rz_t,x,c) = V(z_r,U(r+1z_t,x,c),\ldots,U(tz_t,x,c),U(x,c),\ldots) > V(z_r,U(r+1z_t,y,c),\ldots,U(tz_t,y,c),U(y,c),\ldots) = U(rz_t,y,c).
\]
\(^{40}\)This step would be meaningless if \( U(1c) \) represented how the agent evaluates consumption streams starting at 1 from the perspective of 0, but not necessarily how he evaluates such streams from the perspective of 1. This observation applies to the rest of the proof.
Again, by definition $((z_0, x, w, c') \sim (z_t, y, h, c'))$ means that

\[ V(z_0, U(z_t, x, w), \ldots, U(z_t, x, w, c'), U(x, w, c'), U(w, c'), \ldots) \]
\[ = V(z_0, U(z_t, y, h, c'), \ldots, U(z_t, y, h, c'), U(y, h, c'), U(h, c'), \ldots). \]

Since $U(z_0, x, c) > U(z_0, y, c)$ for all $c$,

\[ V(z_0, U(z_t, y, w, c'), \ldots, U(z_t, y, w, c'), U(y, w, c'), U(w, c'), \ldots) \]
\[ < V(z_0, U(z_t, y, h, c'), \ldots, U(z_t, y, h, c'), U(y, h, c'), U(h, c'), \ldots). \]

This implies that $U(h, c') > U(w, c')$. Otherwise, $U(y, h, c') \leq U(y, w, c')$ and, by induction, $U(s, z_t, y, h, c') \leq U(s, z_t, y, w, c')$ for all $0 \leq s \leq t$, which is a contradiction.

Finally, it must be that $U(x, w, c') > U(y, h, c')$. Otherwise, again by induction, for all $0 \leq s \leq t$

\[ U(s, z_t, y, h, c') > U(s, z_t, x, w, c'), \]

which contradicts $((0, z_t, x, w, c') \sim (0, z_t, y, h, c'))$.

Now suppose that we replace the condition $((0, z_t, x, w, c') \sim (0, z_t, y, h, c'))$ with $(0, z_t, x, t+2z_s, w, c') \sim (0, z_t, y, t+2z_s, h, c')$ where $s \geq t+2$. By the same argument as before, $(0, z_t, y, t+2z_s, w, c') < (0, z_t, y, t+2z_s, h, c')$ and therefore $(h, c') \succ (w, c')$. Otherwise, by induction $(x, t+2z_s, w, c') \gtrsim (x, t+2z_s, h, c')$ for all $0 \leq \tau \leq s$ (where $z_{t+1} = y$). Then, it must be that $(x, t+2z_s, w, c') \succ (y, t+2z_s, h, c')$. Otherwise, since $(x, t+2z_s, h, c') \succ (c, t+2z_s, w, c')$ for all $0 \leq \tau \leq s$, we would have $(0, z_t, x, t+2z_s, h, c') \succ (0, z_t, x, t+2z_s, w, c')$.

12.4 Proof of Theorem 2

This proof adapts arguments in Debreu (1960) and Koopmans [1960, 1964] to the present environment. It is convenient to work in terms of the streams of immediate consumption and future well-being $f$, defined in (4), and the binary relation $\succ^*$ on $\mathcal{F}$ induced by the function $V : \mathcal{F} \to \mathbb{R}$ in the proof of Theorem 1.

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

**Lemma 3.** Axiom 2 implies that $\succ^*$ satisfies the following property. For any $f, f' \in \mathcal{F}$ and $\pi \in \Pi'$,

\[ (f_\pi, f'_\pi) \succ^* (f'_\pi, f_\pi) \iff (f_\pi, f'_\pi) \succ^* (f'_\pi, f_\pi), \]

where $\pi^c = T \setminus \pi$. By Axiom 3, $\succ^*$ depends on $f_0, f_1, f_2$, and $3f$.

**Proof.** Recall that $t c \sim t c'$ implies $U(t c) = U(t c')$, which is equivalent to $f_t = f'_t$. Then, by
Axiom 2, for any $\pi \in \Pi'$

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

By Debreu (1960), there exist then continuous nonconstant functions $\bar{\nabla}, \hat{u}, a, b,$ and $d$ such that

$$\bar{\nabla}(f) = \hat{u}(f_0) + a(f_1) + b(f_2) + d(3f) \quad (19)$$

and

$$f >^* f' \iff \bar{\nabla}(f) > \bar{\nabla}(f').$$

By Lemma 3 with $\pi = \{0\}$, Axiom 8, and Koopmans’s [1960] argument, $V$ can be expressed as

$$V(f) = W(v(f_0), A(f_1)) \quad (20)$$

for some continuous, nonconstant functions $W, v,$ and $A$, where $W$ is strictly increasing. Similarly, by Lemma 3 with $\pi = \{1\}$, Axiom 8, and Koopmans’s [1960] argument, $V$ can be expressed as

$$V(f) = \bar{W}(v(f_0), \bar{A}((\bar{G}(f_1), B(2f))),$$

for some continuous, nonconstant functions $\bar{W}, \bar{A}, \bar{G},$ and $B$, where $\bar{W}$ and $\bar{A}$ are strictly increasing. Now use Axiom 4 to obtain, as shown by Koopmans [1960], that $A$ in (20) and $B$ in (21) are homeomorphic and therefore $B$ can be taken to equal $A$ by a simple modification of the function $\bar{A}$. This leads to

$$V(f) = \hat{W}(v(f_0), \hat{A}((\hat{G}(f_1), A(2f))). \quad (21)$$

According to (20), for every $v(f_0), >^*$ depends on $f_1$ only through $A(1f)$. Therefore, for all $f_1$,

$$A(1f) = \phi_1(a(f_1) + h(2f)) \quad (22)$$

for some strictly increasing and continuous function $\phi_1$, where $h(2f) = b(f_2) + d(3f)$.

According to (21), for every $v(f_0)$ and $\hat{G}(f_1), >^*$ depends on $f_2$ only through $A(2f)$. Therefore, for all $f_2$,

$$A(2f) = \phi_2(h(2f) + d(3f)) \quad (23)$$

for some strictly increasing and continuous function $\phi_2$.

According to (21), for every $v(f_0)$ and $A(2f)$, $>^*$ depends on $f_1$ only through $\hat{G}(f_1)$. Therefore,

$$a(f_1) \equiv G(f_1) = \phi_3(\hat{G}(f_1)), \quad (24)$$
for some strictly increasing and continuous function $\phi_3$.

Now comparing (22) and (23) implies that, for all $f$,

$$a(f_2) + h(3f) = \phi(b(f_2) + d(3f)),$$

where $\phi$ is some strictly increasing continuous function.

**Lemma.** $\phi$ 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 $\phi$ is increasing and continuous. Note that, since $b$, $d$, and $U$ are continuous and non-constant 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 $\overline{b}()$ and $\overline{d}()$ similarly. So

$$\overline{a}(x) + \overline{h}(y) = \phi(\overline{b}(x) + \overline{d}(y) + b(x_0) + d(y_0)) - \phi(b(x_0) + d(y_0))$$

$$\equiv \overline{\phi}(\overline{b}(x) + \overline{d}(y)).$$

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

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

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

Using (24), we can now show that $\overline{\phi}$ is linear, thus $\phi$ is affine. First, note that since $\overline{\phi}(0) = 0$, $\overline{\phi}(z) = -\overline{\phi}(-z)$; so we can focus on the positive part, $\overline{T}_+$, or the negative part, $\overline{T}_-$, of $\overline{T}$. Suppose, without loss of generality, that $\overline{T}_+ \neq \emptyset$. Consider any $b > b' > 0$ in $\overline{T}_+$ such that $b$ and $b'$ are rational. Then, by (24), $\overline{\phi}(b) = m\overline{\phi}(\frac{1}{n})$ and $\overline{\phi}(b') = m'\overline{\phi}(\frac{1}{n'})$ for $m, m', n, n' \in \mathbb{N}$. Since $\overline{\phi}(\frac{1}{n})n = \overline{\phi}(1) = \overline{\phi}(\frac{1}{m'})n'$, it follows that $\overline{\phi}(b) = \frac{b}{b'}\overline{\phi}(b')$. Since rationals are dense in $\overline{T}_+$ and $\overline{\phi}$ is continuous, $\overline{\phi}(b) = \frac{b}{b'}\overline{\phi}(b')$ holds for all $b, b' \in \overline{T}_+$, which implies linearity.

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

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

$$d(3f) = \alpha h(3f) = \alpha(b(f_3) + d(4f)).$$

(25)

It follows that

$$\nabla(f) = \dot{u}(f_0) + G(f_1) + \alpha G(f_2) + d(3f).$$

42
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, $d(f_3) = \hat{d}(f_3')$. By Axiom 4 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 $\varphi$. Again by Axiom 4,
\[
\alpha G(f_2) + \varphi(G(f_3)) \geq \alpha G(f_3') + \varphi(G(f_3'))
\]
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 $\varphi$ must be strictly increasing, it follows that $\alpha < 1$. Finally, by iteratively applying (25) and relying on $\alpha \in (0, 1)$, we get that $\succ^*$ can be represented by
\[
V^*(f) = u(f_0) + \sum_{t=1}^{\infty} \alpha^t G(f_t).
\]
To conclude, we choose, as the utility $\hat{U}$ representing preference $\succ$ over $C$, the function defined by
\[
\hat{U}(c) = V^*(c_0, U(c), \ldots).
\]
The functions $\hat{U}$ and $U$ are strict increasing transformations of one another. Letting $\hat{G}$ denote the function of $\hat{U}$ such that $\hat{G}(\hat{U}(c)) = G(U(c))$ for all $c$, we obtain the representation formula (5). For the uniqueness part, note that the additive form of $\hat{U}$ is unique up to affine transformations, i.e., $\hat{U} = a\hat{U} + b$ for $a > 0$ and $b \in \mathbb{R}$. So,
\[
\hat{U}(c) = au(c_0) + b + \sum_{t=1}^{\infty} \alpha^t aG(\hat{U}(c))
\]
\[
= au(c_0) + b + \sum_{t=1}^{\infty} \alpha^t aG\left(\frac{\hat{U}(c) - b}{a}\right).
\]
Finally, it is easy to see that Axiom 5 holds if and only if $G$ is strictly increasing—using Lemma 2.

12.5 Proof of Proposition 3

(i) To show that $G$ is bounded on $U$, recall from Axiom 7 (Constant-flow Dominance) that $U(c)$ is finite for all $c \in C$. Suppose, by contradiction, that $G$ is unbounded on $U$. Then, for each $n$,
there must be a constant stream \( c^n \) with utility \( U^n \) such that \( G^n = G(U^n) \geq n \), and \( G^n > G^m \) for \( n > m \). Moreover, by Axioms 5, a stream \( c \) that equals \( c^n \) for the first \( k \) periods and \( c^n \) forever after, for \( m > n \), must satisfy \( G(U(c)) \geq n \). This is because \( U(t) > U(c^n) \) for \( t > k \) and hence \( U(kc) > U(ke^n) \); then, by induction, \( U(t) > U(c^n) \) for \( 0 \leq t < k \). Now construct a new stream as follows: for some large \( M \), start with (the consumption defining) \( c^M \) for the first 10 periods, then \( c^{M+10/\alpha} \) for the next 10 periods, then \( c^{M+20/\alpha} \) for the next 10 periods, and so on. By construction, \( U(c) \) must exceed the sum of \( \alpha^t G(U(t)) \) over all \( t \)'s that are multiples of 10 (this is because the remaining terms are nonnegative). But, since by construction \( \alpha^t G(U(t)) \geq M^t \), that sum diverges to infinity and \( U(c) \) must be infinite, violating Axiom 7.

To show that \( U \) is \( H \)-continuous, note that for any \( c, \tilde{c} \in C \)

\[
|U(c) - U(\tilde{c})| = \left| u(c_0) - u(\tilde{c}_0) + \sum_{t=1}^{\infty} \alpha^t [G(U(t)c) - G(U(t)\tilde{c})] \right|
\leq \sum_{t=1}^{T} \alpha^t |G(U(t)c) - G(U(t)\tilde{c})| + \alpha^T \frac{\alpha 2 \tilde{G}}{1 - \alpha},
\]

where \( \tilde{G} = \sup_{\tilde{U} \in \tilde{U}} |G(\tilde{U})| \). So, for any \( \varepsilon > 0 \), choose \( T(\varepsilon) \) so that \( \alpha^{T(\varepsilon)} \frac{\alpha 2 \tilde{G}}{1 - \alpha} < \varepsilon \).

Finally, take \( \nu', \nu \in U \). By definition, there are \( c', c \in C \) such that \( U(c') = \nu' \) and \( U(c) = \nu \). By Lemma 2, we can take \( c' = (x, x, \ldots) \) and \( c = (y, y, \ldots) \) for some \( x, y \in X \). Suppose \( u(x) > u(y) \). Then, \( U(x, y, \ldots) > U(y, y, \ldots) \) and for all finite \( t > 1 \), by induction, \( U(c_0, \ldots, c_t, \ldots) > U(x, y, \ldots) \), where \( c_\tau = x \) for \( 0 < \tau \leq t \). Since \( U \) is \( H \)-continuous, \( U(x, \ldots) > U(y, \ldots) \). By representation (5),

\[
U(x) - \frac{\alpha}{1 - \alpha} G(U(x)) > U(y) - \frac{\alpha}{1 - \alpha} G(U(y)).
\]

Rearranging, we get that for any \( \nu' > \nu \) in \( U \)

\[
\frac{1 - \alpha}{\alpha} (\nu' - \nu) > (G(\nu') - G(\nu)).
\]

(ii) Let \( C(M) \) be the set of consumption streams such that \( |u(c_t)| \leq M \) for all \( t \), and \( B(M) \) be the space of bounded real-valued functions with domain \( C(M) \). Endowed with the sup norm \( \| U \|_{\infty} = \sup_{c \in C(M)} |U(c)| \), \( B(M) \) is a complete metric space. Let \( J \) be the operator on \( B(M) \) defined by

\[
J(U)(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U(t)c).
\]

By construction, \( J(U) \) is bounded over \( C(M) \), as \( u \) is bounded by \( M \) and \( U \) is bounded over \( C(M) \). Moreover, since \( G \) is \( K \)-Lipschitz continuous with \( K < (1 - \alpha)/\alpha \), \( J \) must be a contraction, as is easily checked. So, \( J \) has a unique fixed point; call it \( U_M \). As \( M \) increases, the domain of \( U_M \) increases. However, for any \( M, N \), uniqueness of the fixed point guarantees that \( U_M \) and \( U_N \) coincide on the intersection of their domains. Thus, we obtain a unique solution
Now suppose that where.

Moreover, since show that (a) \( J^{\alpha} < c \) for any \( \{c_n\} \) defined and independent of the chosen sequence. To see this, note that, for any such sequence, time and hence, consumption levels of \( c \) as for any \( c \)

de \( c \) and type constructed above. Let \( U^{\ast} \) be the set of \( H \)-continuous functions. To verify that \( U^{\ast} \) is \( H \)-continuous, it suffices to show that (a) \( J \) maps \( \mathcal{H} \) onto itself, and (b) \( \mathcal{H} \) is closed under the \( \sup \) norm. Indeed, this will guarantee that \( J \)'s fixed-point belongs to \( \mathcal{H} \). To show (a), take any \( U \in \mathcal{H} \) and \( \varepsilon > 0 \). Since \( \alpha < 1 \) and \( G \) is bounded, there is \( T > 0 \) such that \( \frac{\alpha^T G}{1 - \alpha} < \varepsilon / 2 \), where \( G = \sup_{\tilde{U} \in \mathcal{U}} |G(\tilde{U})| \).

Moreover, since \( U \in \mathcal{H} \), there exists \( N \) such \( |U(c) - U(\tilde{c})| < \varepsilon / 2 \) whenever \( c_t = \tilde{c}_t \) for all \( t \leq N \). For any \( c \) and \( \tilde{c} \),

\[
|J(U)(c) - J(U)(\tilde{c})| \leq \left| \sum_{t=1}^{\infty} \alpha^t [G(U(t,c)) - G(U(t,\tilde{c}))] \right| \\
\leq K \sum_{t=1}^{T-1} \alpha^t |U(t,c) - U(t,\tilde{c})| + \alpha^T \frac{2G}{1 - \alpha},
\]

where \( K \) is the Lipschitz constant of \( G \). The first term is less than \( \frac{K\alpha}{1 - \alpha} \max_{t \leq T-1} |U(t,c) - U(t,\tilde{c})| \).

Now suppose that \( c_t = \tilde{c}_t \) for all \( t \leq N' = N + T \). This implies that \( (t,c)_t = (t,\tilde{c})_t \) for all \( t \leq T \) and \( t' \leq N \), because \( t,c \) is truncating at most \( T \) elements of \( c \), and \( c \) and \( \tilde{c} \) were identical up to time \( T + N \), by construction. By definition of \( N \), we have \( |U(t,c) - U(t,\tilde{c})| < \varepsilon / 2 \) for all \( t \leq T \) and, hence, \( |J(U)(c) - J(U)(\tilde{c})| < \varepsilon \). Setting \( T(\varepsilon) = N' \) shows that \( J(U) \) satisfies \( H \)-continuity. To prove (b), consider a sequence \( \{U^n\} \) in \( \mathcal{H} \) that converges to some limit \( U \) in the sup norm. Now fix \( \varepsilon > 0 \). There is \( m \) such that \( \|U^m - U\|_{\infty} < \varepsilon / 3 \). Since \( U^m \in \mathcal{H} \), there is \( N \) such that \( |U^m(c) - U^m(\tilde{c})| < \varepsilon / 3 \) whenever \( c_t = \tilde{c}_t \) for all \( t \leq N \). Thus, for such \( c, \tilde{c} \),

\[
|U(c) - U(\tilde{c})| \leq |U(c) - U^m(c)| + |U^m(c) - U^m(\tilde{c})| + |U^m(\tilde{c}) - U(\tilde{c})| < \varepsilon,
\]

which shows that \( U \in \mathcal{H} \).

To extend the definition of \( U^* \) from \( C(B) \) to \( C \), for any \( c \in C \setminus C(B) \), consider any sequence \( \{c^n\} \) in \( C(B) \) such that \( c^n_t = c_t \) for all \( t \leq n \), and let \( U^*(c) = \lim_{n \to +\infty} U^*(c^n) \). This limit is well-defined and independent of the chosen sequence. To see this, note that, for any such sequence \( \{c^n\} \) and any \( \varepsilon > 0 \), \( H \)-continuity of \( U^* \) implies that there is \( T \) such that \( |U^*(c) - U^*(\tilde{c})| < \varepsilon \) whenever \( c_t = \tilde{c}_t \) for all \( t \leq T \). Hence, \( |U^*(c^n) - U^*(\tilde{c}^n)| < \varepsilon \) for all \( n, m \geq T \), since the consumption levels of \( c^n \) and \( \tilde{c}^m \) coincide up to \( \min\{n, m\} \). So, \( \{U^*(c^n)\} \) forms a Cauchy sequence in \( \mathbb{R} \) and thus converges. Moreover, the limit is independent of the chosen sequence, as for any \( \varepsilon > 0 \), \( |U^*(c^n) - U^*(\tilde{c}^n)| < \varepsilon \) for \( n \) large enough and sequences \( \{c^n\} \) and \( \{\tilde{c}^n\} \) of the type constructed above.

The limit \( U \) thus defined satisfies representation (5). Since \( U^* \) is a fixed point of \( J \) on \( C(B) \) and \( c^n \) belongs to \( C(B) \), for each \( n \)

\[
U^*(c^n) = u(c^n_0) + \sum_{t=1}^{\infty} \alpha^t G(U^*(c^n)),
\]

45
The left-hand side converges to $U^*(c)$. Moreover, for each $t$, $U^*(tc^n)$ converges to $U^*(tc)$ (which is similarly well defined). Since $G$ is continuous, $G(U^*(tc^n))$ converges to $G(U^*(tc^n))$ for each $t$. Since $\alpha < 1$ and $G$ is bounded, by the dominated convergence theorem, the right-hand side converges to $u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U(tc))^\ast$, which proves that (5) holds for all $c \in C$.

Finally, there is a unique $H$-continuous extension of $U^*$ from $C(B)$ to $C$ that solves (5). To see this, let $U$ be any $H$-continuous solution to (5). Since $U$ is a fixed point of $J$ and the fixed point is unique on $C(B)$, $U$ must coincide with $U^*$ on $C(B)$. Take any $c \in C \setminus C(B)$ and $\varepsilon > 0$. By $H$-continuity of $U$ and $U^*$, both $|U(c) - U(\tilde{c})|$ and $|U^*(c) - U^*(\tilde{c})|$ are less than $\varepsilon/2$ for some $\tilde{c} \in C(B)$ equal to $c$ for all $t$ up to a large $N$. Since $U$ and $U^*$ must be equal at $\tilde{c}$, $|U(c) - U^*(c)| < \varepsilon$. Since $\varepsilon$ was arbitrary, $U(c) = U^*(c)$ for all $c$, establishing uniqueness.

### 12.6 Proof of Corollary 1

By Theorem 2, $\succ$ can be represented by

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

Since $(x, c) \succ (y, c)$, $u(x) = u(y) + \pi$ for some $\pi > 0$. Hence, for any $t > 0$, $U(c^x) - U(c^y)$ equals $\pi - \sum_{s=1}^{t} \alpha^s \Delta G_s$, where $\Delta G_s$ is defined recursively as follows: for $s = t$,

$$\Delta G_t = G(U(tc^y)) - G(U(tc^y) - \pi),$$

otherwise

$$\Delta G_s = G(U_s(tc^y)) - G(U_s(tc^y) - \sum_{k=1}^{t-s} \alpha^k \Delta G_{s+k}).$$

By Proposition 3, $\Delta G_t < \frac{1-\alpha}{\alpha} \pi$ and

$$\Delta G_{t-1} = G(U_{t-1}(tc^y)) - G(U_{t-1}(tc^y) - \alpha \Delta G_t)$$

$$< (1 - \alpha) \Delta G_t < \frac{(1 - \alpha)^2}{\alpha} \pi.$$

Now, suppose that, for all $k$ such that $s < k \leq t - 1$, $\Delta G_k < \frac{(1-\alpha)^2}{\alpha} \pi$. It follows that

$$\Delta G_s < \frac{1-\alpha}{\alpha} \left[ \sum_{\tau=1}^{t-s} \alpha^\tau \Delta G_{s+\tau} \right]$$

$$< \frac{1-\alpha}{\alpha} \left[ \sum_{\tau=1}^{t-s-1} \alpha^\tau \frac{(1-\alpha)^2}{\alpha} + \alpha^{t-s} \frac{(1-\alpha)}{\alpha} \right] \pi$$

46
\[
\begin{align*}
&= \frac{(1-\alpha)^2}{\alpha} \left[ \sum_{\tau=0}^{t-s-2} \alpha^\tau (1-\alpha) + \alpha^{t-s-1} \right] \bar{u} \\
&= \frac{(1-\alpha)^2}{\alpha} \bar{u}.
\end{align*}
\]

Therefore,
\[
\sum_{s=1}^{t} \alpha^s \Delta G_s < \pi \left[ \alpha^t \frac{1-\alpha}{\alpha} + \sum_{s=1}^{t-1} \alpha^s \frac{(1-\alpha)^2}{\alpha} \right] = \pi(1-\alpha).
\]

We conclude that \( U(c^x) - U(c^y) > \alpha \bar{u} > 0 \).

12.7 Proof of Corollary 2

By representation (5), \( U \) clearly depends on \( c_0 \) only through \( u_0 = u(c_0) \). This implies that \( U(1c) \) — and hence also \( U(c) \) (from (5)) — depends on \( c_1 \) only through \( u_1 = u(c_1) \). By induction, \( U(c) \) depends on \( (c_0, \ldots, c_t) \) only through \( (u_0, \ldots, u_t) \), for each \( t \). There remains to establish the result at infinity: If \( c \) and \( \hat{c} \) are two streams such that \( u(c_t) = u(\hat{c}_t) \) for all \( t \), we need to show that \( U(c) = U(\hat{c}) \). From the previous step, assume without loss of generality that \( c_t = \hat{c}_t \) for all \( t \leq T \), where \( T \) is any large, finite constant. Since \( U \) is \( H \)-continuous, we can choose \( T \) so that \( |U(c') - U(\hat{c}')| < \varepsilon \) for all \( c', \hat{c}' \) that coincide up to \( T \). Since \( c \) and \( \hat{c} \) satisfy this property, \( |U(c) - U(\hat{c})| < \varepsilon \), and since \( \varepsilon \) was arbitrary, \( U(c) = U(\hat{c}) \). This shows that the sequence \( \{u_t = u(c_t)\}_{t=0}^{\infty} \) of period-utility levels entirely determines the value of \( U(c) \), proving the result.

12.8 Proof of Proposition 4

Consider representation (5) in Theorem 2. For every \( c \in C \), we have sequences \( \{u_s\}_{s=0}^{\infty} \) and \( \{U_s\}_{s=0}^{\infty} \), where \( u_s = u(c_s) \) and \( U_s = U_s(c) \). Since \( u \) is continuous and \( X \) is connected, the range of \( u \) is a connected interval \( I_u \subset \mathbb{R} \). Recall that the range of \( U \) is also a connected interval \( U \subset \mathbb{R} \). Using this notation,
\[
d(t, c) = \frac{\partial U_0/\partial u_t}{\partial U_0/\partial u_0}.
\]

Note that \( \frac{\partial U_s}{\partial u_s} = 1 \) for all \( s \geq 0 \). Since \( G \) is differentiable, we have
\[
\frac{\partial U_0}{\partial u_t} = \sum_{\tau=0}^{t-1} \alpha^{t-\tau} G'(U_{t-\tau}) \frac{\partial U_{t-\tau}}{\partial u_t}.
\]
More generally, for $1 \leq \tau \leq t$,
\[
\frac{\partial U_{t-\tau}}{\partial u_t} = \sum_{s=0}^{\tau-1} \alpha^{\tau-s} G'(U_{t-s}) \frac{\partial U_{t-s}}{\partial u_t}.
\]

So, for $\tau = 1$, $\frac{\partial U_{t-1}}{\partial u_t} = \alpha G'(U_t)$. More generally, for $2 \leq \tau \leq t$,
\[
\frac{\partial U_{t-\tau}}{\partial u_t} = \alpha \sum_{s=0}^{(\tau-1)-1} \alpha^{(\tau-1)-s} G'(U_{t-s}) \frac{\partial U_{t-s}}{\partial u_t} + \alpha G'(U_{t-(\tau-1)}) \frac{\partial U_{t-(\tau-1)}}{\partial u_t}
= \frac{\partial U_{t-(\tau-1)}}{\partial u_t} \alpha(1 + G'(U_{t-(\tau-1)})�).
\]

So,
\[
\frac{\partial U_{t-\tau}}{\partial u_t} = \alpha^\tau G'(U_t) \prod_{s=1}^{\tau-1} (1 + G'(U_{t-s})).
\]

Let $\prod_{s=1}^{\tau-1} (1 + G'(U_{t-s})) = 1$ if $\tau = 1$. Then,
\[
\frac{\partial U_0}{\partial u_t} = \alpha^\tau G'(U_t) + G'(U_t) \sum_{\tau=1}^{t-1} \alpha^\tau G'(U_{t-\tau}) \prod_{s=1}^{\tau-1} (1 + G'(U_{t-s}))
= \alpha^\tau G'(U_t) \left[ 1 + \sum_{\tau=1}^{t-1} G'(U_{t-\tau}) \prod_{s=1}^{\tau-1} (1 + G'(U_{t-s})) \right].
\]

### 12.9 Proof of Theorem 3

Using Axiom 6 and Theorem 2, we also have
\[
(c_0, c_1, 2c) \succ (c_0, c_1, 2c') \iff (c_0, c_1, 2c') \succ (c_0, c_1, 2c') \quad (26)
\]
\[
(c_0, c_1, 2c) \succ (c_0, c_1, 2c') \iff (c_0, c_1, 2c) \succ (c_0, c_1, 2c') \quad (27)
\]
\[
(c_0, c_1, 2c) \succ (c_0, c_1, 2c) \iff (c_0, c_1, 2c) \succ (c_0, c_1, 2c) \quad (28)
\]
\[
(c_0, c_1, 2c) \succ (c_0, c_1, 2c) \iff (c_0, c_1, 2c) \succ (c_0, c_1, 2c) \quad (29)
\]

By Debreu’s [1960], conditions (26)-(29) and (i)-(ii) in Axiom 6 imply that $\succ$ 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 2, $\succ$ is also represented by
\[
u(c_0) + \alpha G(u(c_1) + g(2c)) + \alpha g(2c),
\]

48
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. Since \( G \) must be increasing, without loss of generality let \( G(U) = \gamma U \) with \( \gamma > 0 \). Finally, by Proposition 3, \( \gamma < \frac{1 - \alpha}{\alpha} \).

12.10 Proof of Corollary 6

For every \( c \in C \), consider the sequence \( \{U_s\}_{s=0} \) in the proof of Proposition 4. By Axiom 5, \( c \geq_u c' \) implies \( U_s \geq U'_s \) for all \( s \geq 0 \). It is immediate that, if \( G' \) is increasing (decreasing), then \( d(t, c) \geq (\leq) d(t, c') \) for all \( t > 0 \). On the other hand, suppose \( G' \) is not increasing, i.e., there is \( U > U' \) in \( \mathcal{U} \) such that \( G'(U) < G'(U') \)—the other case is similar. By definition and Lemma 2, \( U = U(c) \) and \( U' = U(c') \) for some constant streams \( c \) and \( c' \). By Axiom 5, \( c \geq_u c' \). However, for all \( t > 0 \), \( d(t, c) < d(t, c') \).

12.11 Proof of Proposition 7

By assumption, for all \( t \),

\[
U(tc) = u(tc) + \sum_{\tau=t+1}^{\infty} \beta\delta^{\tau-t} u(c_{\tau}),
\]

where \( 0 < \beta = \frac{\gamma}{1+\gamma} < 1 \), \( 0 < \delta = (1+\gamma)\alpha < 1 \), \( 0 < \alpha < 1 \).

(\( \Leftarrow \)) See main text.

(\( \Rightarrow \)) Using the expression for \( U \), we get

\[
\sum_{t=0}^{\infty} w(t)U(tc) = w(0)u(c_0) + \sum_{t=1}^{\infty} u(c_t) \left[ w(t) + \beta\delta^{t} \left( \sum_{\tau=0}^{t-1} \frac{w(\tau)}{\delta^\tau} \right) \right].
\]

By assumption, \( \sum_{t=0}^{\infty} w(t)U(tc) = \sum_{t=0}^{\infty} \delta^{t} u(c_t) \). Therefore, the coefficients of \( u(c_t) \) must match for all \( t \). For \( t = 0 \), \( w(0) = 1 \). Then, for \( t = 1 \),

\[
w(1) = (1 - \beta)\delta = \alpha.
\]
Now suppose \( w(t) = \alpha^t \) for all \( t = 0, \ldots, \tau \). Then,

\[
w(\tau + 1) = \delta^{\tau + 1} - \beta \delta^{\tau + 1} \frac{1 - \alpha^{\tau + 1}}{1 - \alpha} = \alpha^{\tau + 1}.
\]

Hence, by induction, \( w(t) = \alpha^t \) for all \( t \).

12.12 Proof of Proposition 6

(Part i) Suppose \( \succ \) can be represented by \( U_t(c) = u(c_t) + \sum_{s \succ t} d(t, s) u(c_s) \). We want to show that there is an alternative representation given by

\[
U_t(c) = u(c_t) + \sum_{s \succ t} q(t, s) U_s(c).
\]

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

\[
u(c_t) = U_t(c_t) - \sum_{s \succ t} q(t, s) U_s(c),
\]

and

\[
U_t(c) = u(c_t) + \sum_{s \succ t} d(t, s) \left[ U_s(c) - \sum_{r \succ s} q(s, r) U_r(c) \right]
\]

\[
= u(c_t) + d(t, t + 1) U_{t+1} + \sum_{s \succ t+1} U_s(c) \left[ d(t, s) - \sum_{t < r \leq s-1} q(r, s) d(t, r) \right]
\]

So, for all \( t \geq 0 \) and \( s > t + 1 \), \( q(t, t + 1) = d(t, t + 1) \) and

\[
q(t, s) = d(t, s) - \sum_{t < r \leq s-1} q(r, s) d(t, r). \tag{30}
\]

Define \( \hat{T}(t, s) \) as in (13). For \( s = t + 2 \), (30) becomes

\[
q(t, t + 2) = d(t, t + 2) - d(t, t + 1) q(t + 1, t + 2)
\]

\[
= d(t, t + 2) - d(t, t + 1) d(t + 1, t + 2).
\]

So (14) holds for all \( t \geq 0 \) and \( s = t + 2 \), since \( \hat{T}(t, t + 2) = \{(t, t + 1, t + 2)\} \). Now suppose that (14) holds for all \( t \geq 0 \) and \( s = t + k \) with \( 2 \leq k \leq n - 1 \). Then, by (30), for \( s' = t + n \)

\[
q(t, s') = d(t, s') - d(t, s' - 1) d(s' - 1, s')
\]

50
\[
- \sum_{t < r \leq s' - 2} d(t, r) \left[ d(r, s') + \sum_{t \in T(r, s')} (-1)^{|t|} \prod_{j=1}^{|t|-1} d(\tau_{j-1}, \tau_j) \right]
\]
\[
= d(t, s') - \sum_{t < r \leq s' - 1} d(r, s') d(t, r) + \sum_{t < r \leq s' - 2} \sum_{t \in T(r, s')} \prod_{j=1}^{|t|-1} d(\tau_{j-1}, \tau_j) d(t, r)
\]
\[
= d(t, s') + \sum_{t \in T(t, s')} (-1)^{|t|} \prod_{j=1}^{|t|-1} d(\tau_{j-1}, \tau_j).
\]

The result follows by induction.

(Part ii) Suppose > can be represented by \( U_t(c) = u(c_t) + \sum_{s > t} q(t, s) U_s(s, c) \). We want to show that there is an alternative representation given by

\[
U_t(c) = u(c_t) + \sum_{s > t} d(t, s) u(c_s)
\]

for some discount function \( d \). If this is true, then (letting \( d(t, t) = 1 \) for all \( t \))

\[
U_t(c) = u(c_t) + \sum_{s > t} q(t, s) U_s(s, c)
\]

\[
= u(c_t) + \sum_{s > t} \left[ \sum_{r \geq s} d(s, r) u(c_r) \right]
\]

\[
= u(c_t) + \sum_{s > t} \left[ \sum_{t < r \leq s} q(t, r) d(r, s) \right] u(c_s).
\]

So, for all \( t \geq 0 \) and \( s > t \),

\[
d(t, s) = \sum_{t < r \leq s} q(t, r) d(r, s).
\]

Therefore, \( d(t, t + 1) = q(t, t + 1) \) for all \( t \geq 0 \). Now define \( T(t, s) \) as in (12). Suppose that (15) holds for all \( t \geq 0 \) and \( s = t + k \) with \( 1 \leq k \leq n - 1 \). Then, by (31)

\[
d(t, t + n) = \sum_{t < r \leq t + n} q(t, r) d(r, s)
\]

\[
= q(t, t + n) + \sum_{t < r \leq t + n - 1} \left[ \sum_{t \in T(t, t + n)} \prod_{j=1}^{|t|-1} q(\tau_{j-1}, \tau_j) q(t, r) \right]
\]

\[
= q(t, t + n) + \sum_{t \in T(t, t + n) \setminus \{(t, t + n)\}} \prod_{j=1}^{|t|-1} q(\tau_{j-1}, \tau_j) q(t, r).
\]
The result follows by induction.

12.13 Proof of Corollary 8

(⇒) The claim follows immediately by iteratively substituting the expression of $U_{t+1}(t+1c)$.

(⇐) The claim follows using condition (30) in the proof of Proposition (6). For any $t \geq 0$, if $s = t + 2$, then

$$q(t, t + 2) = d(t, t + 2) - d(t, t + 1)q(t + 1, t + 2) = 0,$$

using (16) for $d(t, t + 2)$ and $q(t + 1, t + 2) = d(t + 1, t + 2)$. Now take $s = t + n$, for $n > 2$. Suppose $q(t, t + k) = 0$ for all $k = 2, \ldots, n - 1$ and all $t \geq 0$. Then,

$$q(t, t + n) = d(t, t + n) - d(t, t + n - 1)q(t + n - 1, t + n) = 0,$$

again using (16) for $d(t, t + n)$ and $d(t, t + n - 1)$ and $q(t + n - 1, t + n) = d(t + n - 1, t + n)$.

References


