A Theory of Intergenerational Altruism

Simone Galperti
UC, San Diego

Bruno Strulovici*
Northwestern University

January 25, 2017

Abstract

Modeling intergenerational altruism is crucial to evaluate the long-term consequences of current decisions, and requires a set of principles guiding such altruism. We axiomatically develop a theory of pure, direct altruism: Altruism is *pure* if it concerns the total utility (rather than the mere consumption utility) of future generations, and *direct* if it directly incorporates the utility of all future generations. Our axioms deliver a new class of altruistic, forward-looking preferences, whose weight put on the consumption of a future generation generally depends on the consumption of other generations. The only preferences lacking this dependence correspond to the quasi-hyperbolic discounting model, which our theory characterizes. Our approach provides a framework to analyze welfare in the presence of altruistic preferences and addresses technical challenges stemming from the interdependent nature of such preferences.

Keywords: pure altruism, non-paternalistic altruism, time inconsistency, present bias, intergenerational welfare, quasi-hyperbolic discounting, welfare criterion.
JEL Classification: D01, D60, D90

*The authors are grateful to Nageeb S. Ali, Nabil Al-Najjar, James Andreoni, Geir B. Asheim, Doug Bernheim, Eddie Dekel, Jeff Ely, Drew Fudenberg, David Laibson, Bart Lipman, Ehud Kalai, Emir Kamenica, Peter Klibanoff, Mark Machina, Paul Milgrom, David G. Pearce, Todd Sarver, Marciano Siniscalchi, Joel Sobel, and Asher Wolinsky, as well as seminar participants at UC San Diego, Stanford, l'Université de Montréal, RUD 2014, University of Chicago, Yale, NASMES 2014, SITE 2014, UCL, WU St. Louis, UC Berkeley, QSPS Utah, and Pennsylvania State University for useful comments. An early version of this project was developed by the second author 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." A subset of the results appeared in a working paper by the same authors, titled "From Anticipations to Present Bias: A Theory of Forward-Looking Preferences." Strulovici is grateful for financial support from the National Science Foundation (No.1151410) and a Sloan Research Fellowship, as well as for the hospitality of Microsoft Research, New England.

1 Introduction

Intergenerational concerns lie at the heart of many economic decisions: They determine the bequests from parents to their descendants (Bernheim et al. (1985)), which act as a fundamental transmission mechanism in macroeconomic analysis (Barro (1974)); they affect our perception of social mobility and economic inequality (Becker and Tomes (1979)); they play a central role in weighing the environmental consequences of current economic activities, as prominently illustrated by the Stern Review (Stern (2007)) and the ensuing debate on the appropriate level of discounting. Intergenerational altruism also underlies models of capital accumulation (Phelps and Pollak (1968)) and affects the stability of political regimes as already suggested by Tocqueville (1835).[1] By shaping the preferences of households, policymakers, and other economic agents, intergenerational altruism influences how society evaluates and chooses between feasible courses of action.

Although many models of intergenerational altruism have been proposed, an investigation of the fundamental properties of preferences that exhibit altruism towards future generations seems to be missing.[2] Understanding these properties is essential to assess the merits and flaws of the different models available as well as to propose new ones. This paper addresses this issue, exploiting the conceptual and analytical apparatus of modern decision theory to describe and represent altruistic preferences.

We propose an axiomatic theory of pure, direct altruism. A generation exhibits *pure* altruism if it derives utility—besides from its own consumption—from its descendants' overall utility, or "well-being." Pure altruism is sometimes called nonpaternalistic (Ray (1987), Pearce (2008)) as it depends on future generations' overall well-being, not on the specific consumption bundles and other considerations which determine it; in particular, it entails no value judgment about the nature of these components. When future generations are also altruistic, the current generation's pure altruism ascribes to each of them a well-being that includes their own altruism. Pure altruism is commonly assumed in models of intergenerational altruism, such as Barro (1974), Kimball (1987), and Loury (1981).[3] Altruism is said to be *direct* if it directly depends on the well-being of all future

[1] On Tocqueville's hypothesis, see also Lipset (1960).

[2] An exception is the axiomatization in Koopmans (1960), which may be interpreted as a theory of pure *indirect* altruism, as explained in this paper. Like Koopmans (1960), Phelps and Pollak (1968), and a large body of the ensuing literature, this paper focuses on the case of forward-looking preferences in the sense that they do not depend on past generations' consumption. The inclusion of backward-looking preferences raises interesting challenges discussed in Section 6. A large literature on normative social choice, discussed in Section 2, analyzes the implications of axioms such as equity and anonymity to rank consumption streams, which may be compared to the altruistic preferences studied here.

[3] Pure altruism also plays a key role in Bergstrom's (1989) analysis of reciprocal altruism.

generations. For example, environmental policies chosen with a regard towards all future generations reflect direct altruism towards these generations. This property is also common in models of intergenerational altruism (Fels and Zeckhauser (2008), Saez-Marti and Weibull (2005)).[4]

Our analysis takes as primitive the preference $\succ$ of the present generation (hereafter, generation 0) over consumption allocations to itself and future generations.[5] We begin by characterizing the class of generation 0's preferences that have the following representation: For every stream $(c_0, c_1, \ldots)$, where c_t is the consumption of generation $t \geq 0$, generation 0's utility from it can be expressed as

$$U(c_0, c_1, \ldots) = V(c_0, U_1, U_2, \ldots) \tag{1}$$

where $U_t = U(c_t, c_{t+1}, \ldots)$ for $t \geq 1$. We interpret U_t as the well-being that generation 0 assigns to generation t based on the stream $(c_{t,} c_{t+1}, \ldots)$. In particular, these well-beings are all computed with the same function U, which means that generation 0 "projects" its own preference onto future generations to evaluate continuation streams. A motivation for this assumption, beyond tractability, is that generation 0 may not know the preferences of generations who will be born many years into the future and may thus use its own preferences as the best guess. Alternatively, generation 0 may view its preference as embodying some normatively desirable properties which future generations will also adopt. The interpretation of U_t also serves to motivate our axioms and provide insights into the resulting class of models and their properties. If $\succ$ has the utility representation (1), then it is *as if* the current generation exhibited pure altruism. A special case of (1) corresponds to Koopmans' (1960) model of *indirect* pure altruism, in which V depends only on c_0 and U_1. By contrast, direct pure altruism requires that V depend on U_t for *all* $t \geq 1$. Koopmans' model obtains as a limit of direct altruism, as discussed in Section 3.3.

Representation (1) allows for arbitrary relations in generation 0's treatment of distinct future generations. On normative and practical grounds, it is valuable to impose more structure on these relations. We introduce two substantive axioms which, together with

[4] See also Abel (1985) and Pearce (2008). Ray and Wang (2001) analyze a model of direct paternalistic altruism.

[5] The domain of this preference relation consists of infinite consumption streams, as in Koopmans (1960) and Diamond (1965). In this decision-theoretic framework, eliciting a preference relation and testing whether it satisfies some axiom requires a standard observability condition: One must know which of any two streams the current generation would choose if it could implement either of them. In this respect, our approach is identical to these earlier works and a large subsequent literature.

ancillary conditions, narrow the representation down to the additive form

$$U(c_0, c_1, \ldots) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U(c_t, c_{t+1}, \ldots)), \tag{2}$$

where $0 < \alpha < 1$ and the function G is interpreted as the utility that generation 0 derives from future generations' well-being. The axioms imply that G is bounded, strictly increasing, and Lipschitz continuous with modulus $(1-\alpha)/\alpha$, a property which turns out to have important consequences, in particular concerning the continuity and uniqueness of representation (2).

Our first axiom, *intergenerational separability*, adapts to pure altruism a principle which often arises in the literature on impure (or paternalistic) altruism. The axiom rules out any complementarity or substitutability in how the well-being of future generations affects the well-being of generation 0. It thus expresses a form of "neutrality," which the literature has often used in relation to future consumption rather than to future well-being (Koopmans (1960), Asheim (2010)). In a theory of pure altruism, this neutrality naturally concerns future generations' well-being rather than their consumption, since their consumption does not matter *per se* to generation 0. Importantly, however, the axiom—as all axioms of our theory—is formulated in terms of consumption streams, since consumption streams are the primitive objects over which generation 0's preference is defined.

Our second axiom is *altruism stationarity*. It focuses on the altruistic components of generation 0's preference by considering only changes in the consumption of future generations. To illustrate, suppose that $c = (c_0, c_1, \ldots)$ and $c' = (c'_0, c'_1, \ldots)$ give the same consumption to a grandmother ($c_0 = c'_0$) who thinks that her son is indifferent between $(c_1, c_2, \ldots)$ and $(c'_1, c'_2, \ldots)$ after accounting for his consumption and the well-being of his daughter, grandson, and so on. Suppose that the future generations' well-beings induce the grandmother, who cares directly about them, to prefer c to c'. Then, intuitively, altruism stationarity says that if her son were removed from the genealogical sequence—bringing all remaining generations one step closer to the grandmother in the sequence—his mother would continue to prefer $(c_2, c_3, \ldots)$ to $(c'_2, c'_3, \ldots)$ for her granddaughter and so on. From a normative standpoint, the axiom may be viewed as imposing coherence on how generation 0 directly cares about future generations. Altruism stationarity is the axiom, in our theory, which distinguishes direct from indirect pure altruism—hence our model from Koopmans' (1960) model and in particular from the exponentially-discounted-utility model (hereafter, EDU). In Koopmans' (1960) model, if the son is indifferent between $(c_1, c_2, \ldots)$ and $(c'_1, c'_2, \ldots)$, then his mother *must* also be indifferent, given her initial consumption. Once her son is removed from the genealogical sequence, however, the grandmother can strictly prefer $(c_2, c_3, \ldots)$ to $(c'_2, c'_3, \ldots)$ for her granddaughter and so on. As direct altruism becomes arbitrarily negligible, the EDU model arises as a limit of our theory.

Representation (2) has distinctive implications concerning how generation 0 resolves intergenerational trade-offs. First, its well-being can be expressed as a function of the consumption utility of all generations ($u(c_0)$, $u(c_1)$, and so on), which allows us to examine how generation 0 trades off its own consumption utility against that of any future generation t. We find that, even though generation 0 cares directly about the well-being of all future generations, it exhibits positive discounting towards future consumption, in the following sense: If generation 0 can swap consumption utility between itself and some future generation, it always chooses the stream starting with the higher of the two consumption utilities.

Perhaps more surprisingly, generation 0 exhibits *present bias*: Roughly speaking, it is more reluctant to transfer consumption from itself to the next generation, than from any future generation to a later one. To see why, consider a grandmother who cares about her son's well-being (understanding that it depends on his daughter's well-being), but also directly about her granddaughter's well-being. The grandmother can disagree with her son on whether he should transfer more resources to his daughter (say, as a bequest), because they internalize differently the effects of changing her consumption. The son takes into account the daughter's utility through his well-being, while in addition to her son's well-being, the grandmother also cares directly about her granddaughter's well-being. As a result, the grandmother thinks that her son should transfer more resources to his daughter than he wants to. At the same time, however, if she were in her son's position, she would agree on keeping more resources for herself. The grandmother thus appears more willing to delay consumption when it affects future generations than when it directly affects herself. Remarkably, therefore, present-generation bias is not an assumption of our theory, but a logical consequence of direct pure altruism. We will show that this consequences emerges already from representation (1), which does not use the separability and stationarity axioms.

Under representation (2), how generation 0 discounts the consumption utility of some future generation t generally depends on the entire consumption stream it is facing. If the function G is concave, for instance, the consumption utilities of successive generations are substitutes: Increasing the consumption utility of some generation t raises the well-being of earlier generations, thereby reducing the marginal value of later generations' consumption utility. This reduces the effective discount factor applied to these later generations, while the opposite holds if G is convex.

These observations suggest that consumption interdependence disappears if G is linear. This is indeed the case, and with linear G our model takes on a well-known form: quasi-hyperbolic (or β-δ) discounting (Phelps and Pollak (1968)), where β and δ are simple functions of α and the constant slope of G. Linearity of G corresponds to an additional axiom, *consumption independence*, which has two parts: (i) how a grandmother trades off her consumption with that of her son does not depend on the consumption—and hence well-being—of his descendants, and (ii) how she trades off her consumption with that of her son's descendants does not depend on his consumption. A byproduct of our general analysis is thus to provide an axiomatization of Phelps and Pollack's (1968) model of imperfect intergenerational altruism as a model of direct pure altruism. While Phelps and Pollak (1968) viewed the EDU model as capturing perfect intergenerational altruism (p. 186), this paper's perspective suggests that EDU exhibits only indirect pure altruism, offering an opposite view on which of these discounting models—the exponential and the quasi-hyperbolic one—captures a deeper degree of intergenerational altruism.

Direct pure altruism alone has significant implications if we now imagine a sequence of generations, each having the same preference relation—a property commonly known as *time invariance*.[6] If each preference has representation (1) with V strictly increasing in all U's, then the resulting sequence of preferences cannot be time consistent:[7] A course of action for generation t and its descendants that is optimal from the viewpoint of generation t need not be optimal from the viewpoint of generation $s < t$. This conclusion identifies direct pure altruism as a possible source of time inconsistency across generations. Importantly, the model predicts a specific form of inconsistency: By the present-generation bias, each generation tends to disproportionately favor its own satisfaction over that of any future generation.

Time-inconsistent preferences create well-known challenges for welfare analysis as dis-

[6]Time invariance is not needed to characterize the preference of generation 0: The former property compares preferences across multiple generations, while the latter characterization is entirely focused on the preference of a single generation. However, time invariance is a useful additional assumption to assess the time consistency of successive generations whose preferences we have characterized. Besides its analytical convenience, time invariance has additional justifications in our setting: First, under a normative interpretation of our axioms, time invariance captures the idea that all generations are subject to the same normative considerations. Second, it describes a coherence between the interpretation that the current generation projects its own preference onto future generations and the future generations' actual preferences. While the first generation's projection may turn out to be false, assuming it is correct is a useful and natural benchmark.

[7]For this conclusion to hold, it is enough that each generation cares about the well-being of its immediate descendant as well as *some* other future generation.

cussed, among others, by Rubinstein (2003) and Bernheim and Rangel (2007, 2009). This paper makes several contributions to this discussion. As noted earlier, time inconsistency is a necessary consequence of direct pure altruism combined with time invariance. This observation weakens the case for paternalistic interventions, which usually relies on viewing time inconsistency as a form of irrationality. Another implication of our analysis concerns the use of a "libertarian" criterion which measures social welfare with the well-being of generation 0 (that is, $U(c_0, c_1, \ldots)$)—this is what is usually done in the case of Koopmans' (1960) model. Such a criterion seems more appropriate for direct pure altruism—hence time *in*consistency—than for indirect pure altruism—hence time consistency. This assessment is based on the other properties of purely and directly altruistic preferences, such as neutrality and sensitivity towards future generations, whose importance relative to time consistency seems a priori not obvious.[8]

One may argue that, despite generation 0's altruism, a social planner should aggregate the preferences of all generations, assigning an appropriate positive weight to each generation's well-being. We show that pure altruism—whether direct or indirect—makes it difficult to find an aggregator that also renders the planner time consistent.[9] One remarkable exception is representation (2) with linear G, the quasi-hyperbolic model. We show that the standard welfare criterion for β-δ preferences obtained by setting $\beta = 1$ corresponds to an aggregator that weighs the well-being of each generation t (including $t = 0$) by α^t, where α comes from representation (2) and satisfies $\alpha = \delta(1 - \beta)$.[10]

2 Related Literature

The relevance of intergenerational altruism is highlighted by its key role in various economic applications, including optimal national savings (Ramsey (1928), Phelps and Pollak (1968)), economic growth (Bernheim (1989)), charitable giving (Andreoni (1989)), family economics (Bergstrom (1997)), public finance (Barro (1974)), and environmental economics (Weitzman (1999), Dasgupta (2008), Schneider et al. (2012)).

Despite its importance, intergenerational altruism lacks a clear conceptual framework, as emphasized by Ray (1987):

[8]Similar properties have been proposed in the normative social-choice literature as desirable features of a welfare criterion (see, for example, Asheim (2010) and the references therein).

[9]In his study of hedonistic altruism and welfare, Ray (2014) examines welfare criteria that aggregate well-being of altruistic, time-consistent, generations and that are formally similar to those in Section 5 (see also Bernheim (1989)).

[10]For other discussions on the "right" welfare criterion to use for the quasi-hyperbolic discounting model see, for example, O'Donoghue and Rabin (1999), Rubinstein (2003), Bernheim and Rangel (2009).

"the representation of non-paternalistic functions [in terms of total utilities, U] in paternalistic form [in terms of consumption utilities, u] has been the subject of limited attention. A systematic analysis of the relationship between these two frameworks [...] appears to be quite a challenge, especially for models with an infinite horizon (pp. 113–114).

Saez-Marti and Weibull (2005) and Fels and Zeckhauser (2008) derive the mathematical equivalence between the β-δ formula and expression (2) with linear G. Bergstrom (1999) studies systems of utility functions that include altruism towards others, focusing on the infinite regress that they may generate. None of these papers, however, provides an axiomatic foundation of either representation.

Several axiomatizations of the intertemporal preference of a decision maker have been proposed. Koopmans (1960) derives a general model which contains as a special case Samuelson's (1937) EDU model. The difference between Koopmans' analysis and ours is explained when we formally state our axioms. The EDU model obtains as a limit of our theory as direct altruism becomes arbitrarily negligible. Other models that include quasi-hyperbolic discounting have been characterized by Hayashi (2003), Olea and Strzalecki (2014), and Echenique et al. (2016). We compare their approaches to ours in Section 3.3.

Time inconsistency of preferences has been extensively studied since Strotz's (1955) seminal work, and has been given several explanations. In Akerlof (1991), present bias is based on a principle of cognitive psychology which says that decision-makers unduly overweigh relatively more salient or tangible events, such as present consumption relative to future one. In Gul and Pesendorfer (2001), time inconsistency can arise from a general change in the decision-maker's preference over time. Halevy (2008), Saito (2011), and Chakraborty and Halevy (2016) link present bias to the certainty effect. In a setting with uncertainty about the length of the horizon, this literature establishes an equivalence between the decision-maker being disproportionately sensitive to certainty, as in Allais (1953) and Kahneman and Tversky (1979), and exhibiting present bias. In Dasgupta and Maskin (2005) hyperbolic discounting arises from an evolutionary response to the uncertain time of payoff realizations, whereas in Farmer and Geanakoplos (2009) it stems from uncertainty about future discount rates. In Kőszegi and Szeidl (2012), time inconsistency and present bias arise because, when facing a present decision, the agent focuses too much on its immediate consequences, but when considering that same decision ex ante, he is able to focus more on its overall consequences over time. Our theory provides a novel explanation for present bias, as part of a broader conceptual framework of how future consequences of current decision enter into the decision-maker's intertemporal preference.

This paper is also related to the normative social-choice literature studying how to rank intergenerational consumption streams.[11] Its typical approach is to introduce properties that those rankings should satisfy and derive their implications and functional representations. Often suggested properties include sensitivity to and equal treatment of all future generations' well-being. The literature highlights, however, that these properties collide with the standard goal of having a numerical representation of preferences (for example, Basu and Mitra (2003)); the present work considers weaker properties consistent with a numerical representation (see also Asheim et al. (2012)). The social-choice literature also commonly excludes sentiments like altruism from the generations' well-being. From this perspective, one of our contributions is to add this dimension to intergenerational preferences and explore the properties that it implies.

From a methodological perspective, our derivation of representation (2) builds on known results in Debreu (1960) and Koopmans (1960) applied to the sequences of generation 0's consumption and future generations' *well-beings* induced by the consumption streams. A complication arises, however, because the space of such sequences does not have a Cartesian-product structure: Since a generation's well-being depends on its descendants' well-beings, one cannot modify those well-being independently of one another. We propose a way to overcome this difficulty which may be useful in other settings where players' preferences are interrelated. Segal (1992, 1994) provides a framework and general conditions encompassing non-Cartesian spaces to derive representation results. While the interdependence structure arising here makes it challenging to directly verify these conditions, the present analysis suggests an approach to usefully connect such a structure with Segal's framework.

3 Preference Representations

3.1 Preliminaries

Consider an infinite sequence of generations, indexed by $t \in \mathbb{N} = \{0, 1, 2, \ldots\}$. Generation t's consumption is denoted by c_t and belongs to a connected, separable, metric space X; for example, c_t can be a finite vector representing the lifetime consumption of generation t. The streams of consumption starting from generation 0 are denoted by $c = (c_0, c_1, \ldots)$ and

[11] Asheim (2010) provides a detailed review of this literature. Intergenerational welfare has also been extensively discussed in the moral-philosophy literature—for example, to apply to the set of all generations the principles laid out by Rawls (1971).

belong to the set $C = X^{\mathbb{N}}$, which we endow with the sup-norm: $\|c - c'\|_C = \sup_t d(c_t, c'_t)$, where d is a bounded metric on X.[12] For $t \geq 1$, the set of streams starting at t is ${}_tC = X^{\mathbb{N}}$, whose elements will be denoted by ${}_tc = (c_t, c_{t+1}, \ldots)$.

The purpose of this paper is to study the *single* preference of generation 0 over streams of consumption for itself and all future generations, i.e., over C. This preference is denoted by $\succ$, with the symbols $\succsim$ and $\sim$ having the usual meaning, and represents the choices that generation 0 would make for itself and all its descendants. A basic premise is that $\succ$ has a utility representation (as in Koopmans' (1960) Postulate 1). This is ensured by the following axioms.[13]

Axiom 1 (Weak Order)**.** *$\succsim$ is a complete and transitive binary relation.*

Axiom 2 (Continuity)**.** *For all $c \in C$, $\{c' \in C : c' \precsim c\}$ and $\{c' \in C : c' \succsim c\}$ are closed.*

Axiom 3 (Future Constant-Flow Dominance)**.** *For all $c \in C$, there exist $x, y \in X$ such that $(c_0, x, x, \ldots) \precsim c \precsim (c_0, y, y, \ldots)$.*

Axioms 1 and 2 are standard. Axiom 3 captures the following idea: For every stream c, from generation 0's viewpoint there is a sufficiently bad consumption x and a sufficiently good one y such that forcing each future generation to consume x (resp. y) is worse (resp. better) than forcing them to consume according to ${}_1c$. These axioms lead to the following standard result.[14]

Theorem 1 (Utility Representation)**.** *Under axioms 1-3, there exists a continuous function $U : C \to \mathbb{R}$ such that $c \succ c'$ if and only if $U(c) > U(c')$.*

Hereafter, $U(c)$ will represent the *total* utility of generation 0 from stream c. To avoid confusion later with the utility that generation 0 derives from its own consumption, we will use the term *well-being* to refer to $U(c)$.

Since we are interested in studying altruism towards future generations, by assumption the well-being of generation 0 depends on the consumption of later generations. It is also natural that generation 0 cares about its own consumption.

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

[12] Assuming that $d(\cdot,\cdot)$ is bounded is without loss of generality, as we can always replace it by $\hat{d}(\cdot,\cdot) = d(\cdot,\cdot)/(1 + d(\cdot,\cdot))$, which is another metric respecting the initial distances.

[13] This paper assumes, as in Koopmans' (1960, 1964) classic analysis and a large subsequent literature, that generation 0's preference does not depend on past generations' consumption. Relaxing this assumption raises interesting challenges, discussed in Section 6.

[14] The proofs of the main results are in the Appendix A. The remaining proofs are in the Online Appendix.

3.2 Pure-Altruism Representation

Our objective is to investigate which properties of $\succ$ correspond to pure altruism towards future generations. Intuitively, generation 0 exhibits pure altruism if it derives utility from the total utilities—or well-being, in our terminology—of future generations (see, for example, Ray (1987) and Pearce (2008)).[15] Since $\succ$ corresponds to the subjective attitude of generation 0 towards future generations, $\succ$ can only reveal how this generation perceives the well-being that each future generation t derives from the stream ${}_tc$. However, generation 0 may not know future generations' preferences—for instance, for those born many years into the future; or it may have no reason to believe that their preference will systematically differ from its own—for instance, because it views its preference as based on some generally appealing and sensible principles. In these cases, generation 0 may simply "project" its preference onto future generations and use it to assess their well-being. This is the approach that we take in this paper. Other axiomatic models of altruism assume that the decision-maker projects its preference onto its peers (see, for example, Saito (2015)).

This leads to the following general class of representations of purely altruistic preferences. Fixing a representation U of $\succ$, let $\mathcal{U}$ be the range of U and define

$$\mathcal{F} = \{(f_1(c), f_2(c), \ldots) : f_t(c) = U({}_tc) \text{ for } c \in C \text{ and } t \geq 1\}. \tag{3}$$

Note that $\mathcal{F} \subset \mathcal{U}^{\mathbb{N}}$, but $\mathcal{F}$ need not be a Cartesian product because the well-being $U({}_tc)$ ascribed to generation t depends on the well-being ascribed to its descendants. For example, generation 1's maximal well-being—when it is finite—may only be achieved by giving all future generations their highest well-being.

Definition 1 (Pure-Altruism Representation)**.** Preference $\succ$ has a pure-altruism representation if and only if it can be represented by

$$U(c) = V(c_0, U({}_1c), U({}_2c), \ldots) \tag{4}$$

for some function $V : X \times \mathcal{F} \to \mathbb{R}$ that is nonconstant in c_0 and $U({}_tc)$ for some $t \geq 1$.

If generation 0 is purely altruistic, its ranking of any two streams c and c' depends only on its own consumption, c_0 and c'_0, and on how it perceives—through the lens of its current preference—that each future generation t evaluates the continuation streams, ${}_tc$ and ${}_tc'$. Hereafter, we will call $U({}_tc)$ the well-being of generation t, keeping in mind that

[15] By contrast, generation 0 is impurely (or paternalistically) altruistic if it cares only about the utility that other generations derive from what they actually consume. To the extent that generation 0 thinks that future generations are also altruistic, this utility from actual consumption differs from each generation's total utility.

this well-being is evaluated from generation 0's perspective.

Axiom 5 below is the key to obtaining a pure-altruism representation. It captures a minimal property of purely altruistic preferences: Given its own consumption, if generation 0 *thinks* that two consumption streams will render all future generations indifferent, it should also be indifferent.

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

The premise ${}_tc \sim {}_tc'$ captures the thought experiment of generation 0 facing the same continuation stream as generation t will face and using its current preference to compare ${}_tc$ and ${}_tc'$. Axiom 5 rules out the possibility, for instance, that generation 0 prefers c to c' because, despite generating the same stream of current consumption and future well-being, they allocate future consumption differently across generations.

Theorem 2. *Axioms 1–5 hold if and only if* $\succ$ *has a pure-altruism representation.*

Proof. Define $\mathcal{F}_0 = X \times \mathcal{F}$ and let $f_0(c) = c_0$ and $f(c) = (f_0, f_1, f_2, \ldots)$.

($\Rightarrow$) First, we 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 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. Let $(f^*)^{-1}$ denote its inverse and, for any $f \in \mathcal{F}_0$, define

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

By Axiom 5, V is a well-defined function, and $V(f(c)) = U(c)$ for every c. By Axiom 4, V is nonconstant in f_0 and f_t for some $t \geq 1$.

($\Leftarrow$) See Appendix A. $\square$

Axiom 5 requires that generation 0 be indifferent between two streams only if it perceives that *all* future generations—not just generation 1—will be indifferent between the relevant continuations of these streams. Axiom 5 clearly holds if generation 0 cares only, besides its own consumption, about the well-being of generation 1—as in Koopmans' (1960) model. By allowing generation 0's preference to depend on future generations' well-being in a richer way, Axiom 5 is a steppingstone in our approach to modeling intergenerational preferences differently from Koopmans.

[16] In general, there may be several consumption streams in an equivalence class. For a simple illustration, take $U(c) = c_0 + c_1 + c_2 + c_3$, $c = (1, 1, -1, -1, 1, 1, -1, -1, \ldots)$ and $c' = (1, -1, -1, 1, 1, -1, -1, 1, 1, \ldots)$.

To distinguish our approach from the standard one, we introduce the following terminology. As noted, Koopmans' (1960) model satisfies

$$U(c) = V(c_0, U({}_1c));$$

in particular, for EDU we have $U(c) = u(c_0) + \delta U({}_1c)$. In this case, generation 0 is purely altruistic only towards generation 1, and by expecting that all future generations will be altruistic in a similar way, it ends up caring about how c affects all of them but only indirectly through the well-being of generation 1. We will thus say that generation 0 exhibits *indirect pure altruism*.[17] We will say that generation 0 exhibits *direct pure altruism* if its well-being $U(c)$ depends directly on the well-being of all future generations (that is, V in (4) depends on $U({}_tc)$ for all $t \geq 1$).[18] Concretely, a grandmother usually cares about her son's well-being—which depends on his daughter's well-being—and *directly* about her granddaughter's well-being. By contrast, in Koopmans' (1960) model, it is as if the grandmother cares about her granddaughter's well-being only *indirectly* to the extent that it affects her son's well-being.

3.3 Additive Pure-Altruism Representation

To obtain sharper results, we refine the general representation V in (4) by considering preferences that satisfy some form of intergenerational separability and stationarity.

The first axiom captures the idea that $\succ$ is separable across generations—that is, separable between generation 0's consumption and future generations' well-being, as well as across their well-being. Intuitively, this means that the well-being of future generations does not affect how generation 0 enjoys its own consumption. Moreover, the well-being of any group of future generations does not influence how generation 0 trades off the well-being of the remaining generations. As such, this notion of separability captures a form of neutrality of generation 0's altruism towards future generations. To formalize these properties, recall that ${}_tc \sim {}_tc'$ means that $U({}_tc) = U({}_tc')$ and let Π consist of all unions of subsets of $\{\{1\}, \{2\}, \{3, 4, \ldots\}\}$.

Axiom 6 (Intergenerational Separability)**.** *Fix any $\pi \in \Pi$. If c, $\hat{c}$, c', $\hat{c}' \in C$ satisfy*
(i) ${}_tc \sim {}_t\hat{c}$ and ${}_tc' \sim {}_t\hat{c}'$ for all $t \in \pi$,
(ii) ${}_tc \sim {}_tc'$ and ${}_t\hat{c} \sim {}_t\hat{c}'$ for all $t \in \mathbb{N} \setminus \pi$,

[17] An alternative interpretation of the EDU model is that generation 0 does care directly about all future generations, but ignores the fact future generations are altruistic, reducing the well-being of generation t to its consumption utility $u(c_t)$.

[18] This classification is non-exhaustive: Partially direct altruism would entail caring about a subset of future generations, beyond the immediate descendant. This paper focuses on the two extreme cases.

(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 illustrate, set $\pi = \{1, 2\}$ and consider streams c and c' which give the same consumption to generation 0 and yield the same well-being to all future generations, except that generation 1 is better off under c while generation 2 is better off under c'. Suppose, moreover, that generation 0 prefers c to c'. Now, replace c by $\hat{c}$ and c' by $\hat{c}'$, with the new streams chosen so that going from c to $\hat{c}$ and from c' to $\hat{c}'$ leaves the well-being of generations 1 and 2 unchanged, while generation 0's consumption and the well-being of all other generations change by the same amount. According to the axiom, generation 0 should prefer $\hat{c}$ to $\hat{c}'$, thereby continuing to trade off the well-beings of generations 1 and 2 in favor of the former. Axiom 6 is inspired by Debreu's (1960) and Koopmans' (1960) separability axioms, but it differs substantively by requiring that certain continuation streams render generation 0 indifferent (that is, $U({}_tc) = U({}_tc')$), rather that certain physical consumptions be equal (that is, $c_t = c'_t$). While expressed purely in terms of consumption streams, as with earlier theories, the axiom thus captures separability in future generations' well-being rather than in their consumption.

The next axiom imposes a weak monotonicity property of altruistic preferences. First, all else equal, generation 0 is better off if it thinks that generation 1 will be better off. Second, suppose that, for *any* horizon T and *any* common continuation stream c'' after T, generation 0 prefers the consumption of the first T generations implied by stream c to that implied by c'; then, generation 0 also prefers the whole c to c'. Intuitively, this rules out the possibility that the well-being of a generation in the infinite future could overturn how generation 0 ranks c and c', even though, having fixed that generation's well-being, all intermediate generations are better off with c than with c'.

Axiom 7 (Monotonicity)**.** *Let* c *and* c' *be any streams in* C.
(i) If $c_0 = c'_0$, ${}_1c \succ {}_1c'$*, and* ${}_tc \sim {}_tc'$ *for all* $t \geq 2$*, then* $c \succ c'$.
(ii) If for every T *and continuation stream* c'' *we have* $(c_0, c_1, ..., c_T, c'') \succsim (c'_0, c'_1, ..., c'_T, c'')$*, then* $c \succsim c'$.

The axioms introduced so far are all satisfied by Koopmans' (1960) model $U(c) = \hat{V}(u(c_0), U({}_1c))$, where $\hat{V}$ strictly increasing in each argument, hence in particular by the EDU model. They thus define a class of preferences which can exhibit either direct or indirect altruism. It turns out that it is possible to distinguish these kinds of altruism by means of a single additional property, stationarity, which marks the watershed between Koopmans' model and ours. Adding Koopmans' (1960) notion of stationarity to Axioms 1-7 completes the characterization of indirectly altruistic preferences, as explained at the end of this section. The notion of stationarity introduced next takes our theory on a different path which generates direct altruism.[19]

Altruism Stationarity and Direct Pure Altruism

Our notion of stationarity focuses on the altruistic component of generation 0's preference by fixing its own consumption and considering only changes in consumption streams that involve future generations. Intuitively, if generation 0 cares *directly* about the well-being of generations beyond its immediate descendant in a coherent way, it should be possible to "remove" generation 1 and preserve how generation 0 ranks the consumption streams starting from generation 2 onward. Clearly, for this comparison to be meaningful, we must start from a situation in which how generation 0 ranks two streams in the presence of generation 1 depends only on the well-being of subsequent generations but not on that of generation 1.

Axiom 8 (Altruism Stationarity)**.** *If* $c, c' \in C$ *satisfy* $c_0 = c'_0$ *and* ${}_1c \sim {}_1c'$*, then* $c \succsim c'$ *if and only if* $(c_0, {}_2c) \succsim (c'_0, {}_2c')$*.*

To gain intuition, suppose that c and c' give the same consumption to a grandmother and that she thinks that her son is indifferent between ${}_1c$ and ${}_1c'$ after he takes into account his consumption as well as the well-being of his daughter and later descendants. Also, suppose that later descendants' well-beings induce the grandmother, who cares directly about them, to prefer c to c'. The axiom says that if the son were to die—which moves all remaining generations one step closer in the genealogical sequence—his mother should continue to prefer ${}_2c$ to ${}_2c'$ for her granddaughter and later descendants.

These axioms lead to the representation in Theorem 3 below. Intuitively, by the theorem it is *as if* generation 0 derives a utility u from its consumption, as usual, and an altruism utility G from the well-being of each future generation, which it discounts exponentially as generations move farther away in the lineage.[20] Moreover, G is bounded, so future generations' well-being can have only a limited impact on generation 0's well-being. Intuitively, this generation cannot become infinitely happy or unhappy just from its altruism towards its descendants. The consumption utility u, however, can be unbounded.

[19] As noted, indirect altruism arises at the limit, as direct altruism becomes negligible. This point is particularly clear for the quasi-hyperbolic discounting case discussed in Section 4.3.

[20] Rogers (1994) suggests an evolutionary justification based on genetic relationship for why generation 0 may progressively care less about its children, grandchildren, and so on: Every step in the lineage reduces the share of genes that generation 0 can expect to have in common with that generation.

Theorem 3 (Additive Pure-Altruism Representation)**.** *Axioms 1-8 hold if and only if the function U may be chosen so that*

$$U(c) = u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U({}_tc)) \tag{5}$$

where $\alpha \in (0,1)$, $u : X \to \mathbb{R}$ and $G : \mathcal{U} \to \mathbb{R}$ are continuous, nonconstant functions, and G is strictly increasing and bounded. Moreover, if (u, α, G) and $(\hat{u}, \hat{\alpha}, \hat{G})$ both represent $\succ$, then $\hat{\alpha} = \alpha$ and there exist $a > 0$ and $b \in \mathbb{R}$ such that $\hat{u}(x) = au(x) + b$ and $\hat{G}(\hat{U}) = aG(\frac{\hat{U}-b}{a})$ for all x and $\hat{U}$.

The proof relies on known results in Debreu (1960) and Koopmans (1960). However, a complication arises in our setting. As noted, the set $\mathcal{F}$ of streams of future generations' well-being induced by consumption streams in C may not be a Cartesian product (see (3)). Roughly speaking, to deal with this issue, the key is to show that (i) if we take any stream f in $\mathcal{F}$, there is an open neighborhood of f which belongs to $\mathcal{F}$ and has the structure of a Cartesian product, and (ii) it is possible to "cover" $\mathcal{F}$ with countably many of such neighborhoods which intersect with each other. Given (i) and (ii), we can obtain a preliminary additive representation on each neighborhood. Relying on these representations' uniqueness up to positive affine transformations, we can then "glue" all of them into a single representation over the entire set $\mathcal{F}$.[21]

To work with this class of models, it is natural to ask whether expression (5) is well defined for any function G. Proposition 1, below, provides necessary and sufficient restrictions for (5) to be well defined. It also proves that the representation U is continuous at infinity in the following sense: The effect on generation 0's well-being of future consumption changes becomes arbitrarily small when these changes take place arbitrarily far in the future.

Definition 2. A function $U : C \to \mathbb{R}$ is H-continuous[22] if, for every $\varepsilon > 0$, there exists $T \in \mathbb{N}$ such that $|U(c) - U(\tilde{c})| < \varepsilon$ whenever $c_t = \tilde{c}_t$ for $t \leq T$.

Proposition 1.
(i) In representation (5), U is H-continuous and for ν', $\nu \in \mathcal{U}$

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

[21] This approach may be useful more generally to obtain separable models of a decision-maker who exhibits pure altruism towards other individuals in society and takes into account that they are also purely altruistic towards each other.

[22] This notion is similar to the concept of "continuity at infinity" of payoff functions in infinite-horizon games (see, for example, Fudenberg and Tirole (1991)).

(ii) Suppose that G is strictly increasing, bounded, and Lipschitz continuous with constant $K < \frac{1-\alpha}{\alpha}$. Then, there exists a unique H-continuous function $U : C \to \mathbb{R}$ that solves (5).

This result helps to choose G appropriately in applications, and implies several properties of directly altruistic preferences, to be described shortly.

Koopmans' Stationarity and Indirect Pure Altruism

Koopmans' (1960) model is based on the general representation $U(c) = \hat{V}(u(c_0), U({}_1c))$ with $\hat{V}$ strictly increasing in each argument. To obtain such a representation in the framework of the present paper, it is enough to add Koopmans' stationarity axiom to Axioms 1-7.[23]

Axiom 9 (Koopmans' Stationarity)**.** *For some c_0 and all ${}_1c$, ${}_1c'$,*

$$(c_0, {}_1c) \succsim (c_0, {}_1c') \text{ if and only if } {}_1c \succsim {}_1c'.$$

As Koopmans (1960) showed, if we want to specialize $\hat{V}$ to obtain the EDU model, we need stronger separability assumptions (cf. Section 4.3).

To see why Axiom 9 rules out direct pure altruism, note that Axiom 6 with $\pi = \{1, 2, \ldots\}$ implies that, if the condition of Axiom 9 holds from some c_0, then it holds for all $c_0 \in X$. Axiom 9 thus implies that, given her consumption c_0, how a grandmother ranks the consequences of streams ${}_1c$ and ${}_1c'$ for her son, granddaughter, great granddaughter, and so on is always pinned down by how she thinks that her son ranks these consequences for himself and his descendants. This property seems unduly restrictive and unrealistic: Why, when assessing the consequences of its actions on future generations, should a generation rely exclusively on its immediate descendant's preferences? This exclusive reliance arises as a limit of our model, when direct pure altruism becomes negligible.

4 Intergenerational Trade-offs

This section studies the consequences of direct pure altruism on how a generation trades off consumption between generations. A special case of representation (5) corresponds to the β-δ discounting model introduced by Phelps and Pollak (1968) to capture "imperfect" intergenerational altruism. Our theory provides a very different perspective, suggesting that the altruism captured by the β-δ discounting model is pure and direct.

[23]This set of axioms imply postulates 1-4 in Koopmans (1960).

4.1 Positive Discounting and Present-generation Bias

Pure direct altruism implies that a generation derives utility from the well-being of future generations and recognizes that the well-being of each future generation depends on its descendants' well-being. Can this induce a generation to sacrifice its own consumption utility in favor of that of a future generation? In the case of representation (5), the answer is negative: Generation 0 always discounts positively its descendants' consumption utility, relative to its own. To formalize this, we introduce a precise definition of positive discounting.

Definition 3 (Positive Discounting)**.** $\succ$ is said to exhibit positive discounting if the following holds: For any x, $y \in X$ such that $(x, c) \succ (y, c)$ for all $c \in C$, and any $t \geq 1$, we have $c^x \succ c^y$ whenever $c_0^x = x$, $c_t^x = y$, $c_0^y = y$, $c_t^y = x$, and $c_s^x = c_s^y$ for all $s \in \mathbb{N} \setminus \{0, t\}$.

The first part of the condition says that x is unambiguously preferred to y. The second part says that, given any two streams which differ only at times 0 and t, with the streams swapping x and y at times 0 and t, the stream starting with x is always preferred to the stream starting with y. The necessary restrictions derived in Proposition 1, above, yield the following key result.

Corollary 1. *If axioms 1-8 hold, then $\succ$ exhibits positive discounting.*[24]

Positive discounting refers to how generation 0 trades off its own consumption utility with future generations', but how does it trade off satisfaction between future generations? Does it do so in the same way in which it trades off its own satisfaction against that of any of those generations? As it turns out, pure direct altruism alone—without imposing stationarity or separability—introduces a systematic difference between trade-offs which directly involve generation 0 and trade-offs which involve only future generations. Perhaps more surprisingly, direct pure altruism leads generation 0 to be biased in favor of its present satisfaction in the following sense.

Definition 4 (Present-generation bias)**.** $\succ$ exhibits present-generation bias if the following holds: For any x, y such that $(x, c) \succ (y, c)$ for all c, and streams such that $(z_0, \ldots, z_{t-1}, x, \hat{x}, c') \sim (z_0, \ldots, z_{t-1}, y, \hat{y}, c')$ for some $t \geq 1$, we have $(x, \hat{x}, c') \succ (y, \hat{y}, c')$.

[24] In a finite-horizon setting, a representation of the form in (5) is consistent with negative discounting: Generation 0 may be willing to sacrifice its satisfaction for the benefit of future generations that are sufficiently close to it in the lineage. Unlike the infinite-horizon case, (5) yields well-defined preferences even if the function G is "steeper" than $\frac{1-\alpha}{\alpha}$. See the earlier version of this paper (Galperti and Strulovici (2014)).

Intuitively, suppose that x is unambiguously preferred to y and that generation 0 is indifferent if generation t is forced to give up x in exchange for y, provided that the change is "compensated" by giving $\hat{y}$ to generation $t+1$ instead of $\hat{x}$. Then, if generation 0 itself faced the choice between x and y in the present, it would strictly prefer *not* to give up x for y even if generation 1 got $\hat{y}$ in exchange.

As the next result shows, this form of present bias is already implied by pure direct altruism, without requiring the separability and stationarity axioms used to characterize representation (5).

Proposition 2. *Suppose that $U(c) = V(c_0, U({}_1c), U({}_2c), \ldots)$ represent $\succ$ and that V is strictly increasing in $U({}_tc)$ for all $t \geq 1$. Then, $\succ$ exhibits present-generation bias.*

The proof of the proposition establishes a stronger form of present bias: The conclusion of Definition 4 holds when $\hat{x}$ and $\hat{y}$ concern any period $s \geq t+1$, instead of just $t+1$.

The logic of Proposition 2 may be explained as follows. Set $t = 1$ in Definition 4 and imagine that x and y correspond to higher and lower levels of consumption. From the viewpoint of a grandmother (generation 0), a reduction in her son's consumption from x to y can be compensated by an appropriate increase in her granddaughter's consumption, $\hat{y}$. However, because the grandmother takes directly into account her granddaughter's well-being, the level $\hat{y}$ that renders her indifferent may be insufficient to render her son indifferent. This implies that if the grandmother were in the same situation of her son, she would also strictly prefer not to reduce her own consumption from x to y even if her child received $\hat{y}$. Thus, the grandmother appears to be more willing to shift consumption to a later generation when only future generations are affected than when she is directly affected.

Proposition 2 generalizes and offers an interpretation to the results obtained by Saez-Marti and Weibull (2005) and Fels and Zeckhauser (2008). These papers assume that generation 0's preferences are represented by an aggregator V which is additive and linear in future generations's well-beings, U. For this specific case, they show that $U(c)$ can be rewritten as the discounted sum of future generations' consumption utilities—given by some function $u(c_t)$—where discounting takes the quasi-hyperbolic form as in Phelps and Pollak (1968) (see Corollary 3 below). This, of course, implies that the preference corresponding to such a $U(c)$ exhibits present-generation bias. Proposition 2 shows that this conclusion holds more generally as a consequence of direct pure altruism.

4.2 Intergenerational Rate of Utility Substitution

Representation (5) exhibits separability in future generations' *well-being* but not necessarily in their *consumption.* In fact, how generation 0 trades off its consumption and that of generation t generally depends on the consumption of all other generations, because generation t's consumption affects the well-being of all preceding generations, and in a way that depends on the consumption of all subsequent generations.

To examine this, we consider generation 0's discount function between 0 and t. Clearly, intergenerational trade-offs involving consumption also depend on the curvature of the consumption utility u. To bypass this dependence, we start with a convenient result, which follows from Theorem 3.

Corollary 2 (u-Representation)**.** *Given representation* (5)*, there exists a nonconstant function* $\hat{U} : \mathcal{I}_u^{\mathbb{N}} \to \mathbb{R}$ *(where* $\mathcal{I}_u$ *is* u*'s range) such that, for all* $c \in C$*,*

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

This implies that, given a stream c, what matters for generation 0's preferences is the streams of consumption *utility* $u_s = u(c_s)$, rather than consumption per se. This simplification is useful because utility levels are real numbers, whereas consumption bundles take values in a potentially much more complex space. The discount function can now be defined as

$$d(t, c) = \frac{\partial \hat{U}(u_0, u_1, \ldots)/\partial u_t}{\partial \hat{U}(u_0, u_1, \ldots)/\partial u_0}. \tag{6}$$

That is, $d(t, c)$ is the marginal rate at which generation 0 substitutes consumption utility between 0 and t. In the EDU model, $d(t, c) = \delta^t$. For $d(t, c)$ to be well defined, the derivatives in (6) must exist. This is always the case when G is differentiable.[25]

Proposition 3. *Suppose that* G *is differentiable. Then,* $d(1, c) = \alpha G'(U({}_1c))$ *and, for* $t \geq 2$*,*

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

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

To understand this formula, suppose that $u(c_t)$ rises by a small amount. This has two effects: (i) generation t's well-being rises, which explains the term $G'(U({}_tc))$; consequently, (ii) the well-being of each generation $\tau \in [1, t]$ also rises, which explains the

[25]Differentiability is a mild additional assumption, since G's monotonicity already implies that it is differentiable almost everywhere.

summation. Moreover, the rise in $U({}_t c)$ affects $U({}_{t-\tau} c)$ through the well-being of all generations between $t - \tau$ and t, which explains the product.

For an arbitrary altruism function G, the discount function $d(t, c)$ thus depends not only on generation t's consumption, but on the entire stream c. Suppose, to illustrate, that G is concave. A drop in the consumption of some future generation s reduces the well-being of earlier generations, which increases the G' terms and, hence, the discount factor $d(t, c)$: Generation 0 becomes more willing sacrifice its current satisfaction $u(c_0)$ in favor of future generations'. The only case in which $d(t, c)$ is independent of c is when G is linear, in which case, the discount factor takes a well-known form.[26]

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

$$d(t, c) = \beta \delta^t,$$

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

Proof. By Proposition 3, the result is immediate for $t = 1, 2$. For $t \geq 3$,

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

□

Thus the class of preferences that Phelps and Pollak (1968) invented to model "imperfectly" altruistic generations is actually equivalent to a specific version of preferences that exhibit direct pure altruism. By contrast, the EDU model—which Phelps and Pollak (1968) viewed as modeling "perfectly" altruistic generations—is actually a specific version of preferences that exhibit only indirect pure altruism. The present paper reverses the common view concerning which of the exponential and β-δ discounting models captures a deeper degree of intergenerational altruism.

The EDU model obtains at the limit of ours as direct altruism towards generations $t \geq 2$ becomes negligible. This limit is obtained by taking α to zero while keeping the product $\alpha\gamma$ equal to some constant $\hat{\delta}$—which necessitates that γ becomes arbitrarily large. The weight that generation 0 puts on generation t is $\gamma\alpha^t$, which converges to zero except for $t = 1$. In this case, Corollary 3 implies that the preference has a beta-delta representation, where $\beta = \gamma/(1+\gamma)$ converges to 1 and $\delta = (1+\gamma)\alpha$ converges to $\hat{\delta} = \gamma\alpha$, which corresponds to the EDU model.

[26] A formally equivalent result appears in Saez-Marti and Weibull (2005) and Fels and Zeckhauser (2008).

4.3 Quasi-hyperbolic Discounting of Consumption Utilities

Since the linearity of G implies quasi-hyperbolic discounting, we wish to determine which properties of the preference imply this linearity. When G is nonlinear, we observed earlier that generation 0's trade-off between its consumption and that of generation t typically depends on the well-being of other generations. Building on this observation, Axiom 10 rules out this type of dependence, and yields the desired conclusion.

Axiom 10 (Consumption 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 a grandmother trades off her consumption with that of her son in a way that does not depend on the consumption and, hence, the well-being of her son's descendants. Condition (ii) says that she trades off her consumption with that of her son's descendants in a way that does not depend on her son's consumption.

Theorem 4 (Linear Pure-Altruism Representation)**.** *Axiom 1-8 and 10 hold if and only if the function U may be chosen so that*

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

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 1-8 and 10 hold if and only if there exist β, $\delta \in (0, 1)$ and a continuous nonconstant function $u : X \to \mathbb{R}$ such that*

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

Proof. By Theorem 4, $U(c)$ is a strictly increasing, linear function of the well-beings $U({}_tc)$ and of $u(c_0)$. By induction , $U(c)$ is a strictly increasing, linear function of the consumption utilities $u(c_t)$ for $t \geq 0$. There exists then a function $\kappa(t) : \mathbb{N} \setminus \{0\} \to \mathbb{R}_{++}$ such that

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

Clearly, for all $t \geq 1$, $\kappa(t) = d(t, c)$ defined in (6). Corollary 3 then shows the claim. □

This result allows us to understand β-δ discounting of consumption utilities in terms of simple properties of generation 0's directly altruistic preference. This preference depends on two, conceptually different, entities: generation 0's physical consumption and the overall well-being of all future generations. This sheds light on why generation 0 treats its consumption utility differently from the *consumption utilities* of all future generations, in a way that is, however, coherent across them as captured by altruism stationarity (Axiom 8). Moreover, by expecting all future generations to also be altruistic, generation 0 develops the view that future generations should be more willing to sacrifice their consumption for the good of their descendants than how much it itself is willing to sacrifice its own consumption for them (present-generation bias). These properties contribute to explaining why the different treatment of future consumption utilities takes a stark form in Phelps and Pollack's (1968) model, where those utilities are all scaled uniformly by $\beta < 1$. Finally, generation 0 evaluates its consumption and future generations' well-being in a separable way, and trades off its consumption against that of any future generation independently of other generations' well-being. This delivers additivity in terms of consumption utilities. Note that Axiom 10 coincides with Koopmans' (1960) strong separability axiom (Postulate $3'$), which in his model delivers EDU. Therefore, according to the present paper, one single axiom—the stationarity one—distinguishes exponential and quasi-hyperbolic discounting. Put differently, in terms of the consumption-utility representation (4), Koopmans' stationarity corresponds to setting $\beta = 1$, whereas altruism stationarity corresponds to setting $\beta < 1$.

Corollary 4 provides a tight link between the degree of present bias, β, and the marginal utility from altruism, γ. In this linear representation, it is possible to interpret γ as the degree to which generation 0 finds future generations' well-being "imaginable" or "vivid." Note that this is distinct from how much generation 0 cares about its descendants as the intergenerational gap increases, which is captured by α^t. Thus, we identify two possible determinants of how much generation 0 is altruistic towards generation t. Corollary 4 implies that β is directly proportional to γ, so generation 0 becomes less present biased when it anticipates more vividly the well-being of future generations.[27]

Other Axiomatizations of β-δ Discounting

Other papers axiomatize preferences over consumption streams that correspond to the β-δ model (Hayashi (2003); Olea and Strzalecki (2014)). These papers take the perspective of a single individual instead of different generations.[28] To attempt a comparison with the present paper, note that generation 0—the decision-maker in our framework—

[27] Vividness of the well-being of future generations implied by today's decisions may be influenced with specific information campaigns. For example, consider the dramatic images that media and environmental organizations report on the catastrophic consequences of manmade climate change.

[28] While Phelps and Pollak (1968) focused on an intergenerational interpretation of quasi-hyperbolic discounting, the single-individual interpretation is equally widespread (following Laibson (1997)).

essentially treats its descendants as future copies of itself. Thus, from a single-individual perspective, one may view our analysis as describing an agent who perceives himself as a collection of selves, one for each period t, and exhibits direct pure altruism towards his future selves, expecting that they will do the same. Therefore, self 0 directly cares about the well-being (that is, $U({}_tc)$) of all his future selves through what we may call *intra*personal altruism,[29] an attitude that is revealed by his current preference over consumption streams.

The property that self 0 directly cares about all his future selves' *well-being* is the first, key conceptual difference from the previous axiomatizations of β-δ discounting. Both Hayashi (2003) and Olea and Strzalecki (2014) adopt the common view that the decision-maker cares only about his per-period consumption utilities.[30] Within this framework, Koopmans' (1960) stationarity (Axiom 9) is replaced by a weaker notion, quasi-stationarity, which means stationarity from period 1 onward. However, if the decision-maker evaluates streams based only on consumption utilities and views consumption in the same way in all periods, it is unclear why stationarity should hold between period 1 and later periods, but not between period 0 and period 1. This issue does not arise with altruism stationarity (Axiom 8), as self 1's well-being is equivalent to the well-being of all future selves, but is different from self 0's physical consumption.

Second, to obtain the β-δ representation, Olea and Strzalecki (2014) need to ensure that current and future per-period utilities are cardinally equivalent. Their axioms lead to novel, practical experiments to identify and measure β and δ, but may be difficult to interpret. That paper imposes an additional, explicit present-bias axiom to obtain $\beta \leq 1$. By contrast, when self-0 directly cares about the well-being of all his future, similarly altruistic selves, present bias follows as a general, logical consequence (Proposition 2). In our theory, β-δ discounting is tightly linked with an intuitive consumption-independence condition (Axiom 10).

Echenique et al. (2016) propose a different method to characterize several models of

[29]The idea that an individual over time consists of a collection of selves is not new in economics (see, for example, Strotz (1955) who thought that "the individual over time is an infinity of individuals" (p. 179) and Frederick (2003)); it also has a long tradition in philosophy (see, for example, Parfit (1971, 1976, 1982)). In this case, direct pure altruism simply means that the individual's present self cares about his future selves and anticipates that they will continue to do the same.

[30]The second paper, in particular, is based on the idea of annuity compensation: To avoid relying on assumptions of the form of u to elicit β and δ separately, the idea is to consider fix compensation levels—and hence fixed u's—and vary the time horizon at which they occur, so as to find exact points of indifference for the decision-maker and hence infer the parameters of the model. The decision-maker has different subjective views of the time distance between period 0 and 1 and between any two future periods. If he cares only about the u's that he gets in each period, then it is possible to objectively space out these u's in an appropriate way so as to identify β and δ.

intertemporal preferences, including the EDU and β-δ model. Their starting point is a dataset consisting of a decision-maker's choices of finite consumption streams from standard intertemporal budget sets and the dated prices that define them. They discuss versions of the Generalized Axiom of Revealed Preference that satisfy various separability requirements adapted to intertemporal settings. Using real data from the experiment in Andreoni and Sprenger (2012), they apply their axioms to classify subjects across models: Roughly, only one third of the 97 subjects is consistent with either EDU or β-δ discounting, and about half violates time separability in consumption.

4.4 Time Inconsistency

The analysis thus far has concerned the preference of a single generation. This section considers the implications of direct pure altruism on how *different* generations evaluate a common consumption stream.

It is well known that if each generation has the same preference and this preference can be represented by Koopmans' (1960) recursive model, then as a whole these preferences form a sequence which satisfies time consistency. That is, if a course of action starting at time t is preferable according to generation t's preference, then it remains preferable *for time t onward* according to generation $t-1$'s preference. To formalize this, consider the family $\{\succ^t\}_{t=0}^{\infty}$, where $\succ^t$ is the preference relation of generation t.[31]

Definition 5 (Time Consistency)**.** $\{\succ^t\}_{t=0}^{\infty}$ exhibits time consistency if, for every $t \geq 0$, $s \geq 1$, and $c, c', \hat{c} \in C$, the following condition holds:

$$(\hat{c}_t, \ldots, \hat{c}_{t+s-1}, c) \succsim^t (\hat{c}_t, \ldots \hat{c}_{t+s-1}, c') \text{ if and only if } c \succsim^{t+s} c'.$$

Despite its similarity with Axiom 9 when $s = 1$, time consistency is conceptually different because it involves multiple, possibly different, preference relations.[32] Indeed, Koopmans et al. (1964) write,

> "[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

[31] Definition 5 expresses in formal terms the idea of time consistency informally introduced by Strotz (1955) in his seminal paper (p. 171).

[32] This point has also been recently emphasized in the experimental analysis of Halevy (2015). It is straightforward to construct examples of preferences $\{\succ^t\}_{t=0}^{\infty}$ that satisfy Axiom 9 but are not time consistent, and vice versa.

of preferences as the time of choice changes is therefore extraneous to the present study." (Koopmans et al. (1964), p. 85, emphasis in the original)

Nonetheless, under the assumption that $\{\succ^t\}_{t=0}^{\infty}$ is time invariant—that is, there exists a common preference relation $\hat{\succ}$ such that $\succ^t = \hat{\succ}$ for all $t \geq 0$—time consistency is equivalent to Koopmans' (1960) stationarity. As noted, this notion of stationarity implies that each generation exhibits indirect altruism. The converse of this is that direct altruism is incompatible with time consistency, a statement proved in the Supplemental Material.

Proposition 4. *Consider a family of preferences $\{\succ^t\}_{t=0}^{\infty}$ and suppose that $\succ^t = \succ^0$ for all $t \geq 1$ and that $\succ^0$ has a pure-altruism representation. Then, $\{\succ^t\}_{t=0}^{\infty}$ 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.

Proposition 4 identifies and highlights a possible source of time inconsistency across generations. This source corresponds to the intuitive property that each generation directly cares about the well-being of future generations beyond its immediate descendant—as in our running example of a grandmother, her son, and her granddaughter. It is common to view time consistency of intergenerational preferences as the norm and time inconsistency as an exception. Proposition 4 reverses this view. If we deem plausible that each altruistic generation cares about future generations' well-being beyond its immediate descendant and expects them to do the same, then we have to conclude that time inconsistency should be the norm, not the exception.[33]

In general, time inconsistency can take many forms. For instance, if we allow the preferences to violate time invariance and differ in arbitrary ways across generations, we could have that generation t prefers higher consumption (say, of fossil fuels) for itself and all future generations, whereas generation $t+1$ prefers lower consumption for itself and all subsequent generations. This kind of inconsistency is clearly possible, but differs from the one arising here. In particular, directly altruistic preferences are time inconsistent

[33] To see the intuition behind Proposition 4, suppose that ${}_1c$ and ${}_1c'$ represent two courses of actions that start in period 1 and involve different consumptions only for generations 1 and 2 (that is, $c_1 \neq c_1'$, $c_2 \neq c_2'$, but ${}_3c = {}_3c'$). In this case, generation 0 (the grandmother) and generation 1 (her son) may disagree on the ranking of ${}_1c$ and ${}_1c'$ for the following reason. The son trades off how his well-being varies because his current consumption changes from c_1 to c_1' and because his daughter's well-being is affected by the change from c_2 to c_2'. In addition to these two effects on her son's well-being, the grandmother also takes *directly* into account how her granddaughter's well-being varies between c_2 and c_2'. Therefore, the grandmother and her son internalize the effects of changing c_2 in different ways, which can induce them to disagree on which course of action is better.

even if *all* generations have the same preference relation. Moreover, by Proposition 2 this inconsistency takes a very specific form: Each generation tends to disproportionately favor its own satisfaction over that of any future generation.

4.5 Intergenerational Trade-offs with Backward-looking Altruism

While our theory's focus on forward-looking agents is standard, incorporating backward-looking altruism to the theory would be instructive and realistic. An axiomatic treatment of two-sided altruism raises conceptual and technical challenges beyond the scope of the present paper—some of which is discussed in Section 6 along with the related literature. As a first step in this direction, this section explores how the intergenerational trade-offs discussed earlier would be affected by the addition of backward-looking altruism. As it turns out, backward-looking altruism has rather counter-intuitive consequences.

We illustrate these consequences in settings with two and three generations. Suppose, first, that there are only two generations, the second of which may be affected by the first generation's well-being. Specifically, consider the linear specification

$$U_0 = u_0 + \alpha U_1 \qquad U_1 = u_1 + \hat{\alpha} U_0$$

where u_i is generation i's utility from consumption and α and $\hat{\alpha}$ are nonnegative factors used to discount future and past generations' well-being. This system has a unique solution as long as the product $\alpha\hat{\alpha}$ is strictly less than 1, yielding

$$U_0 = \frac{1}{1-\alpha\hat{\alpha}}(u_0 + \alpha u_1) \qquad U_1 = \frac{1}{1-\alpha\hat{\alpha}}(u_1 + \hat{\alpha} u_0).$$

Ignoring the scaling factor $1/(1-\alpha\hat{\alpha})$, notice that the weight put by generation 0 on generation 1's consumption utility, relative to its own, is equal to α and thus independent of generation 1's altruism coefficient $\hat{\alpha}$. Remarkably, therefore, adding backward-looking altruism does not affect how generation 0 trades off the consumption utilities of both generations. This result has the following intuition. With backward-looking altruism, increasing 0's consumption utility now raises 0's well-being by increasing 1's well-being, which 0 finds valuable. But increasing 1's consumption utility increases 1's well-being directly, and both increases, whether coming from 1's backward-looking altruism or from a direct increase of u_1, get injected into 0's well-being in the same way through 0's forward-looking altruism. Therefore, consumption utilities u_0 and u_1 have the same relative impact on 0's well-being with this new channel as they did over the original forward-looking channel.

The effect of backward-looking altruism gets even more surprising—in fact, counter-intuitive—if one adds a third generation. Suppose that

$$\begin{aligned} U_0 &= u_0 + \alpha U_1 + \alpha^2 U_2 \\ U_1 &= u_1 + \alpha U_2 + \hat{\alpha} U_0 \\ U_2 &= u_2 + \hat{\alpha} U_1 + \hat{\alpha}^2 U_0. \end{aligned}$$

This system has a unique solution provided $\alpha\hat{\alpha} < \frac{1}{3}$, yielding

$$U_0 = \frac{1}{(1+\alpha\hat{\alpha})(1-3\alpha\hat{\alpha})}[(1-\alpha\hat{\alpha})u_0 + \alpha(1+\alpha\hat{\alpha})u_1 + 2\alpha^2 u_2].$$

Now the weights of u_1 and u_2 have *increased*, relative u_0's weight, as a result of generation 1's and 2's altruism towards older generations ($\hat{\alpha} > 0$).[34] This is counterintuitive as one would expect that generation 0's consumption matters more now that it affects generation 1's and 2's well-being. To gain intuition, it is useful to compare this situation with the two-generation case, where the relative weights on 0's and 1's consumption utility on 0's well-being were independent of $\hat{\alpha}$. Consider the increase of u_2's weight relative to u_0's weight. Backward-looking altruism causes generation 0's well-being to affect the well-being of generations 1 and 2. By forward-looking altruism, u_2 also affects the well-being of generation 1 and 2 by affecting generation 0's well-being. Therefore, if these were the only effects of u_2, its weight relative to u_0 would not depend on the amount of backward-looking altruism, just as in the two-generation case. In addition, however, now there is an intermediate generation between 0 and 2. Therefore, u_2 also enters the backward-looking channel that links 1's well-being to 2's well-being; and this additional effect—from which u_0 is excluded—increases the impact of u_2 relative to u_0. The increase in u_1's relative weight has a similar intuition.

5 Welfare Analysis with Intergenerational Altruism

Models featuring time-inconsistent preferences raise serious conceptual problems when defining welfare criteria and discuss policy questions (see, for example, Rubinstein (2003); Bernheim and Rangel (2007, 2009)). Concerning β-δ discounting, Rubinstein (2003) notes that

> "Policy questions were freely discussed in these models even though welfare assessment is particularly tricky in the presence of time inconsistency. The literature

[34] Note that the weight of u_1 relative to u_0, $\alpha(1+\alpha\hat{\alpha})/(1-\alpha\hat{\alpha})$, as well as the weight of u_2 relative to u_0, $2\alpha^2/(1-\alpha\hat{\alpha})$, are increasing in $\hat{\alpha}$.

often assumed, though with some hesitation, that the welfare criterion is the utility function with stationary discounting rate δ (which is independent of β)." (p. 1208)

A more fundamental question is whether time inconsistency across generations' preferences justifies some form of paternalistic social planner. An immediate consequence of the present paper is to *weaken* the case for paternalistic interventions. If intergenerational time inconsistency were the result of some form of bounded rationality or lack of consideration for future generations, one may be tempted to argue that society benefits from delegating its choices to a paternalistic planner. This argument, however, is moot if time inconsistency is the logical consequence of direct pure altruism which every generation fully takes into account. In this case, the current generation 0 is already taking into account the preferences that it expects all future generations to have. Thus, unless the social planner believes that generation 0 anticipates these preferences in a systematically biased way, one may argue that she should simply adopt a "libertarian" stance and measure the welfare of any stream c using $W^L(c) = U(c)$.

This libertarian criterion coincides with the criterion that is usually applied for standard time-consistent models, such as EDU, but has significantly different implications in the present model. For standard models, W^L corresponds to evaluating consumption streams for the entire society based only on the consumption of generation 0 and how it thinks that generation 1 will rank the remaining entire streams. In our model, by contrast, W^L evaluates streams based on the consumption of generation 0 and how it thinks that *all* future generations will rank the streams they will face for their remaining future. For example, consider two policies, A and B, inducing streams c^A and c^B such that ${}_1c^A \sim {}_1c^B$ but ${}_tc^A \succ {}_tc^B$ for all $t \geq 2$. Then, the criterion W^L based on the standard model implies that policy A is as desirable as B. By contrast, the criterion W^L based on our model implies that A is strictly more desirable. Thus, the second criterion is more sensitive to a policy long-run consequences for all future generations' well-being. From this perspective, it is far from obvious that such a criterion should be dismissed simply because it violates time consistency.

Despite generation 0's direct pure altruism, one may still argue that a libertarian criterion does not sufficiently take into account the preferences of future generations and attempt to include these preferences above and beyond how they affect generation 0. This leads to a paternalistic welfare criterion aggregating the preferences (that is, well-being) across all generations (Bernheim (1989), Phelan (2006), Farhi and Werning (2007)). Crucially, one may ask whether there exists an aggregator which, despite the preferences' time inconsistency, renders the social planner *time consistent*. For the additive criterion

$$W^P(c) = \sum_{t=0}^{\infty} w(t) U({}_tc), \tag{7}$$

where $w : \mathbb{N} \to \mathbb{R}_{++}$, we obtain the following result.

Proposition 1. *Suppose that all generations have the same preference representations, given by Theorem 4 and Corollary 4 with parameters (u, α, γ) and (u, β, δ), respectively. Then,*

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

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

This proposition provides a novel justification for the welfare criterion, consisting in dropping the β factor, cited by Rubinstein. We focus here on proving the "if" direction, which has an intuitive justification, relegating the proof of the "only if" direction to Appendix A. Let $W(c) = \sum_{t=0}^{\infty} \alpha^t U({}_tc)$. Using the linear representation of $U(c)$ in Theorem 4, we have

$$W(c) = u(c_0) + (1+\gamma)\alpha \sum_{t \geq 1} \alpha^{t-1} U({}_tc) = u(c_0) + (1+\gamma)\alpha W({}_1c).$$

Comparing the extremes of this equation already shows why the planner must be time consistent. It also shows that $W(c)$ must equal the discounted sum of consumption utilities with discount factor $\hat{\delta} = (1+\gamma)\alpha$. We know from Corollaries 3 and 4 that $\hat{\delta} = \delta$ in the β-δ version of $U(c)$, which proves the "if" part. There is a more intuitive argument, however. Note that the linear representation satisfies

$$U(c) = u(c_0) + \gamma\alpha W({}_1c) = u(c_0) + \frac{\gamma\alpha}{\hat{\delta}} \sum_{t \geq 1} \hat{\delta}^t u(c_t).$$

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

Proposition 1 implies, for the widespread Phelps–Pollak model of intergenerational altruism, the existence of a paternalistic welfare criterion that aggregates the time-inconsistent preferences of all generations, yet renders the planner time consistent. This criterion is completely determined by generation 0's revealed preference, summarized by u, α, and γ and is easily interpreted by viewing α as a survival probability used by both the social planner and each successive generation to assess the relevance of each successive period. However, the conclusion of Proposition 1 does not hold for more general directly altruistic preferences.[35]

[35] For a general function G, the criterion $W^P(c) = \sum_{t=0}^{\infty} \alpha^t U({}_tc)$ need not satisfy time consistency. It is possible to construct examples in which $\sum_{t=1}^{\infty} \alpha^t G(U({}_tc)) > \sum_{t=1}^{\infty} \alpha^t G(U({}_tc'))$ and $\sum_{t=1}^{\infty} \alpha^t [G(U({}_tc)) +$

By contrast, aggregating time-consistent preferences by a similar process leads to a time-*inconsistent* welfare criterion. In particular, using the EDU model $U(c) = \sum_{t=0}^{\infty} \delta^t u(c_t)$, we may rely on δ to aggregate well-being across generations using the criterion $\hat{W}(c) = \sum_{t=0}^{\infty} \delta^t U({}_tc)$ (cf. Ray (2014)). Doing so makes the planner time inconsistent, since it can be shown that $\hat{W}(c) = u(c_0) + \sum_{t=1}^{\infty} \delta^t (1+t) u(c_t)$. Except for Phelps and Pollack's (1968) quasi-hyperbolic model, pure altruism across generations—whether direct or indirect—thus makes it difficult to construct a welfare criterion that includes the preferences of all generations and renders a planner time consistent.

6 Concluding Remarks

We have examined how pure direct altruism towards future generations shapes the preference of a generation over courses of action that have long-lasting consequences. Each generation exhibits positive discounting and a present-generation bias. The dependence of each generation's well-being on that of its descendants gives rise to rich discounting patterns, the simplest of which turns out to be quasi-hyperbolic discounting. Our analysis characterizes directly altruistic preferences within a classic decision-theoretic framework and in terms of axioms that are easy to interpret. Given the importance of uncertainty for long-run problems—such as poverty or climate change—future work should extend the theory to allow for stochastic consumption streams.

A natural question is whether the theory in this paper is descriptive or normative. We view these interpretations as non-exclusive and both valid and provide rigorous conceptual tools which may be used for both. On the one hand, our axioms are meant to capture general properties that one might expect to describe how each generation ranks consumption streams for itself and its descendants. Our results then identify which utility functions satisfy these properties and help understand their key behavioral implications. On the other hand, one can view our axioms as crystallizing some ethical principles—whose origin lies outside the model—that all generations adopt and share as guides for intergenerational altruism. Such principles might be that one should be sensitive to all of one's descendants' preferences (to the extent that they may be anticipated) and should try to treat those preferences in a neutral way. For instance, the ongoing public debate on the control of greenhouse-gas emissions often stresses the importance of taking into account the consequences of these decisions for many generations far into

$U({}_tc)] > \sum_{t=1}^{\infty} \alpha^t [G(U({}_tc')) + U({}_tc')]$, but $\sum_{t=1}^{\infty} \alpha^t U({}_tc) < \sum_{t=1}^{\infty} \alpha^t U({}_tc')$. In this case, $W^P(c_0, {}_1c) > W^P(c_0, {}_1c')$, but $W^P({}_1c) < W^P({}_1c')$.

the future, but usually does not consider complementarities or complicated dependence across these generations. These principles may provide a normative foundation for our main axioms—intergenerational separability and altruism stationarity—and hence for our main representation theorem. More generally, the paper points out that requiring preferences to be time consistent on the basis of some normative consideration comes at the cost of violating other properties which may also have a normative appeal.

Our welfare analysis also lends itself to positive and normative interpretations. On the one hand, it connects generations' altruistic preferences with the welfare criterion used by a social planner. This is important if, for instance, the planner coincides with a democratically elected government which responds only to the preference of the current generation: Its altruism will be the only channel through which the government will take into account future generations. In this case, the paper emphasizes that the government will inherit the current generation's desire for long-run sustainable policies (such as adopting renewable sources of energy) but also its present bias, which may render commitments to those policies desirable. On the other hand, the aggregating rules that we considered may again be based on some external ethical principles—for example, that the planner should directly weigh the well-being of each generation in a way that includes its own altruism. These principles, *combined* with direct pure altruism, may give a normative meaning to the time-consistent criterion that we derive in our welfare analysis. Our results highlight that, at least for a well-known and broadly-used model of intertemporal preferences, it is possible to reconcile time inconsistency at the generations' level with time consistency at the planner's level.

This paper proposes a theory of *forward-looking* altruism, which is consistent with a large literature on intergenerational preferences and instrumental in comparing our approach with other theories of forward-looking agents. The addition of backward-looking altruism can affect intergenerational trade-offs in surprising and even counter-intuitive ways, as we saw in Section 4.5. Beyond this, incorporating backward-looking altruism into an axiomatic approach raises a number of conceptual challenges.[36] One is an observability problem: how can we determine how a generation's preference depends on the consumption of its ancestors, when this consumption is past and thus unchangeable? A potential resolution is to extract from a generation's preference information about how this generation perceives that its descendants will care about itself, so as to back out the backward-looking component of altruism. Another challenge concerns the infinite

[36]Backward-looking altruism is considered by Peleg and Yaari (1973), Pearce (2008), Kimball (1987), Bergstrom (1989), and Bergstrom (1999), who assume specific utility functions and derive conditions leading to well-defined preferences and equilibrium existence.

regress stemming from reciprocal altruism, and the technical difficulties that it raises for an axiomatic formulation. Addressing the reciprocity challenge would pave the way to an axiomatic theory of altruism across contemporaneous members of society, extending the analysis beyond backward-looking altruism. Incorporating backward-looking and contemporaneous altruism into an axiomatic approach is an exciting avenue for future research; the framework and techniques developed in this paper should prove helpful to this end.[37]

[37] In particular, a well-known problem when writing a reciprocal altruism model is that the utility function may "blow up." However, if one starts with preferences as primitives, then the class of model emerging from axioms capturing reciprocal altruism will come with specific parametric restrictions—also implied by the axioms—guaranteeing that the models of the class are well specified and do not blow up. Even in our theory of forward-looking altruism, the altruism utility G must be $(1-\alpha)/\alpha$-Lipschitz continuous, by Proposition 1, which guarantees that the discounting effect dominates the forward-looking-altruism effect so that utility is always well-defined and finite in the additive representation (5).

A Appendix: Proofs of the Main Results

A.1 Proof of the "only if" part of Theorem 2

Suppose that there exists a continuous function $U : C \to \mathbb{R}$ which represents $\succ$ and that $U(c) = V(f(c))$ where $V : \mathcal{F}_0 \to \mathbb{R}$ is a function which is nonconstant in f_0 and f_t for some $t \geq 1$. This last property immediately implies that $\succ$ satisfies Axiom 4. Since U is a continuous representation of $\succsim$, standard arguments imply that this relation satisfies Axioms 1-3. Finally, consider any streams c, $c' \in C$ that satisfy $c_0 = c'_0$ and ${}_tc \sim {}_tc'$ for all $t \geq 1$. Then, since U represent $\succ$, we have $U({}_tc) = U({}_tc')$ for all $t \geq 1$; that is, $f_t(c) = f_t(c')$ for all $t \geq 1$. Since by assumption $f_0(c) = c_0 = c'_0 = f_0(c')$ and V is a function, we must have $V(f(c)) = V(f(c'))$ or, in other words, $U(c) = U(c')$, which in turn implies that $c \sim c'$ as required by Axiom 5.

A.2 Proof of Theorem 3

By the definition of $\mathcal{F}$ in (3), note that $\mathcal{F}$ need not be a Cartesian product, as f_t depends on f_s for $s > t$. Letting ${}_tf = (f_t, f_{t+1}, \ldots)$, we can denote elements in $\mathcal{F}$ by $(f_1, f_2, \ldots, f_{t-1}, {}_tf)$. On $\mathcal{F}_0 = X \times \mathcal{F}$ (where $f_0(c) = c_0$), the primitive $\succ$ induces a $\succ^*$ with representation $V : \mathcal{F}_0 \to \mathbb{R}$; by Theorem 2, $\succ^*$ is well defined. If $\mathcal{F}_0$ were a Cartesian product, we could mimic the steps in Debreu (1960) (Theorem 3) and Koopmans (1972) on the domain $\mathcal{F}_0$ to prove our theorem. However, this is not possible. We will then proceed as follows. In step 1, we show that $\succ^*$ satisfies the essentiality and strong separability properties at the heart of Debreu's (1960) Theorem 3. In step 2, we show that the ranking of streams $(f_0, f_1, f_2, {}_3f) \in \mathcal{F}_0$ depends only on a function of ${}_3f$; so we can restrict attention to a four dimensional space. In step 3, we show that this space is a Cartesian product 'locally;' so we can apply Debreu's result to obtain an additive representation 'locally.' Since additive representations are unique up to positive, affine transformations, we can extend uniquely the additive representation to the entire $\mathcal{F}_0$. In step 4, we show that this representation takes the form in our Theorem 3.

Step 1. Lemma 1 says that, if $(f_0, f_1, f_2, {}_3f) \succ^* (f'_0, f'_1, f'_2, {}_3f')$, then changing the common components of $(f_0, f_1, f_2, {}_3f)$ and $(f'_0, f'_1, f'_2, {}_3f')$ in the same way leaves the ranking under $\succ^*$ unchanged.

Lemma 1. *Fix any nonempty subset π of $\{0, 1, 2, 3\}$. Then*

$$(f_0, f_1, f_2, {}_3f) \succ^* (f'_0, f'_1, f'_2, {}_3f') \Leftrightarrow (\hat{f}_0, \hat{f}_1, \hat{f}_2, {}_3\hat{f}) \succ^* (\hat{f}'_0, \hat{f}'_1, \hat{f}'_2, {}_3\hat{f}'),$$

where $f_t = \hat{f}_t$, $f'_t = \hat{f}'_t$, ${}_3f = {}_3\hat{f}$, and ${}_3f' = {}_3\hat{f}'$ if t or 3 are in π, and $f_t = f'_t$, $\hat{f}_t = \hat{f}'_t$, ${}_3f = {}_3f'$, and ${}_3\hat{f} = {}_3\hat{f}'$ if t or 3 are not in π.

Proof. Recall that ${}_tc \sim {}_tc'$ is equivalent to $f_t = f'_t$. Then, by Axiom 6, for any π

$$V(f_0, f_1, f_2, {}_3f) > V(f'_0, f'_1, f'_2, {}_3f') \Leftrightarrow V(\hat{f}_0, \hat{f}_1, \hat{f}_2, {}_3\hat{f}) > V(\hat{f}'_0, \hat{f}'_1, \hat{f}'_2, {}_3\hat{f}').$$

$\square$

Using Lemma 1 with $\pi = \{0\}$ and $\pi = \{1, 2, 3\}$, we obtain the following.

Lemma 2. *The function $V : \mathcal{F}_0 \to \mathbb{R}$ can be written in the form*

$$V(f) = W(u(f_0), d({}_1 f)), \tag{8}$$

where $u : X \to \mathcal{I}_u \subset \mathbb{R}$ and $d : \mathcal{F} \to \mathcal{D} \subset \mathbb{R}$. W is jointly continuous in its two arguments and strictly increasing in each of them, u is continuous, and $\mathcal{I}_u$ and $\mathcal{D}$ are non-degenerate intervals.

Proof. Consider $\succ^*$ on $X \times \mathcal{F}$ and Lemma 1 with $\pi = \{0\}$ and $\pi = \{1, 2, 3\}$. By an argument similar to that in Section 5 of Koopmans' (1960), for any f we can write $V(f) = W(u(f_0), d({}_1 f))$, where $u(f_0) = V(f_0, {}_1\hat{f})$ for some ${}_1\hat{f} \in \mathcal{F}$ and $d({}_1 f) = V(f'_0, {}_1 f)$ for some $f'_0 \in X$. Recall that $V(f(c)) = U(c)$ for all $c \in C$. Hence, the continuity property of U implies continuity of u. By Axiom 4, neither u nor d can be constant. Since X is connected, u takes all values in a connected interval $\mathcal{I}_u \subset \mathbb{R}$. Since $d({}_1 f(c)) = U(f'_0, {}_1 c)$, U is continuous, and X is connected, d takes all values in a connected interval $\mathcal{D} \subset \mathbb{R}$. By definition of u and Lemma 1 with $\pi = \{0\}$, W must be strictly increasing in its first argument on $\mathcal{I}_u$. Similarly, by definition of d and Lemma 1 with $\pi = \{1, 2, 3\}$, W must be strictly increasing in its second argument on $\mathcal{D}$. Given $\hat{c}$, $U(\cdot, {}_1\hat{c})$ takes values in an interval. Then the strictly increasing $W(\cdot, d({}_1 f(\hat{c})))$ also takes values in an interval and hence must be continuous in its first argument on $\mathcal{I}_u$. By a similar argument, W must be continuous in its second argument, and hence jointly continuous on $\mathcal{I}_u \times \mathcal{D}$.

□

Hereafter, let $\overline{u} = \sup \mathcal{I}_u$ and $\underline{u} = \inf \mathcal{I}_u$. Also note that the function d in Lemma 2 defines a ranking on $\mathcal{F}$.

Lemma 3. *There exist $x, y, z, x', y', z' \in X$ and $c \in C$ such that (i) $(z, c) \succ (z', c)$, (ii) $(y, z, c) \sim (y'z', c)$, and (iii) $(x, y, z, c) \sim (x', y', z', c)$.*

Proof. By Axiom 4, there exist z, $z' \in X$ and $c \in C$ such that $(z, c) \succ (z', c)$. Using representation (8), we have $u(z) > u(z')$. Now, consider (y', z, c) and (y, z', c) where $y = z$ and $y' = z'$. By Axiom 7(i), $(y', z, c) \succ (y', z', c)$.

Case 1: $(y, z', c) \succsim (y', z, c)$. Since $\mathcal{I}_u$ is connected, we can modify y to $y'' \in X$ so that $u(y'')$ takes any value in $[u(y'), u(y)]$. By Axiom 2, there exists y'' such that $(y'', z', c) \sim (y', z, c)$; moreover, we must have $u(y'') > u(y')$. Now consider (x, y'', z', c) and (x', y', z, c) where $x = z$ and $x' = z'$. By Axiom 7(i), $(x', z, c) \succ (x', z', c)$; so, by Axiom 8, $(x', y', z, c) \succ (x', y'', z', c)$.

Case 1.1: $(x, y'', z', c) \succsim (x', y', z, c)$. Since we can modify x to $x'' \in X$ so that $u(x'')$ takes any value in $[u(x'), u(x)]$, by Axiom 2, there exists x'' such that $(x'', y'', z', c) \sim (x', y', z, c)$.

Case 1.2: $(x, y'', z', c) \prec (x', y', z, c)$. We can modify z and y'' to $\tilde{y}, \tilde{z} \in X$ so that $u(\tilde{z})$ and $u(\tilde{y})$ take any value in $[u(z'), u(z)]$ and $[u(y'), u(y'')]$. Moreover, we can do so maintaining $(\tilde{y}, z', c) \sim (y', \tilde{z}, c)$ by Axiom 2. Since $(x, y', z', c) \succ (x', y', z', c)$, by Axiom 2, there exist $\tilde{y}$ and $\tilde{z}$ such that $(x, \tilde{y}, z', c) \sim (x', y', \tilde{z}, c)$. Finally, we must have $u(\tilde{z}) > u(z')$, so $(\tilde{z}, c) \succ (z', c)$.

Case 2: $(y, z', c) \prec (y', z, c)$. We can modify z to $\hat{z} \in X$ so that $u(\hat{z})$ takes any value in $[u(z'), u(z)]$. Since $(y, z', c) \succ (y', z', c)$, by Axiom 2 there exists $\hat{z}$ such that $(y, z', c) \sim (y', \hat{z}, c)$, and by Axiom 7(i) we must have $(\hat{z}, c) \succ (z', c)$ and hence $u(\hat{z}) > u(z')$. Now consider (x, y, z', c) and $(x', y', \hat{z}, c)$ where $x = z$ and $x' = z'$.

Case 2.1: $(x, y, z', c) \succsim (x', y', \hat{z}, c)$. We can modify x to $\hat{x}$ so that $u(\hat{x})$ takes any value in $[u(x'), u(x)]$. By Axiom 7(i), $(x', z', c) \prec (x', \hat{z}, c)$; so, by Axiom 8, $(x', y, z', c) \prec (x', y', \hat{z}, c)$. Then, by Axiom 2 there exists $\hat{x}$ such that $(\hat{x}, y, z', c) \sim (x', y', \hat{z}, c)$.

Case 2.2: $(x, y, z', c) \prec (x', y', \hat{z}, c)$. We can modify y and $\hat{z}$ to $\hat{y}$ and $\hat{z}'$ so that $u(\hat{y})$ and $u(\hat{z}')$ take any value in $[u(y'), u(y)]$ and $[u(z'), u(\hat{z})]$. Moreover, we can do so maintaining $(\hat{y}, z', c) \sim (y', \hat{z}', c)$ by Axiom 2. Since $(x, y', z', c) \succ (x', y', z', c)$, by Axiom 2 there exists $\hat{y}$ and $\hat{z}'$ such that $(x, \hat{y}, z', c) \sim (x', y', \hat{z}', c)$.

□

Hereafter, for $t \in \{0, 1, 2, 3\}$, we will refer to the factor t of $\mathcal{F}_0$ as the component of position $t + 1$ in the representation $(f_0, f_1, f_2, {}_3f)$ of every $f \in \mathcal{F}_0$ (e.g., the factor 2 is the third component of every $(f_0, f_1, f_2, {}_3f) \in \mathcal{F}_0$).

Definition 6 (Debreu (1960)). For $t \in \{0, 1, 2\}$, if $f \succ^* f'$ for some $f, f' \in \mathcal{F}_0$ with $f_s = f'_s$ for all $s \neq t$, then the factor t of $\mathcal{F}_0$ is called essential for $\succ^*$. If $(f_0, f_1, f_2, {}_3f) \succ^* (f'_0, f'_1, f'_2, {}_3f')$ for some $f, f' \in \mathcal{F}_0$ with $f_s = f'_s$ for $s = 0, 1, 2$, then the factor 3 is called essential for $\succ^*$.

Lemma 4. *For all $t \in \{0, 1, 2, 3\}$, the factor t of $\mathcal{F}_0$ is essential.*

Proof. By Axiom 4, the factor 0 is essential. Using the streams in Lemma 3, let ${}_1c = (x, y, z, c)$ and ${}_1c' = (x', y', z', c)$ and consider the corresponding f and f' in $\mathcal{F}_0$ with any $f_0 = f'_0$. We have $f_1 = f'_1$, $f_2 = f'_2$, $f_3 > f'_3$, and $f_t = f'_t$ for all $t \geq 4$. By Axiom 7(i), $(f_0, f_3, f_4, {}_5f) \succ^* (f'_0, f'_3, f'_4, {}_5f')$, hence the factor 1 is essential. By Axiom 8, $(f_0, f_2, f_3, {}_4f) \succ^* (f'_0, f'_2, f'_3, {}_4f')$ and $(f_0, f_1, f_2, {}_3f) \succ^* (f'_0, f'_1, f'_2, {}_3f')$. So the factors 2 and 3 are essential.

□

Step 2. By Lemma 1 with $\pi = \{2, 3\}$, $\succ^*$ also satisfies the following property:

$$(f_0, f_1, {}_2f) \succ^* (f_0, f_1, {}_2f') \Leftrightarrow (\hat{f}_0, \hat{f}_1, {}_2f) \succ^* (\hat{f}_0, \hat{f}_1, {}_2f').$$

Define $\tilde{\mathcal{Q}} = \{(f_1(c), d({}_2f(c))) : c \in C\}$. Note that $\tilde{\mathcal{Q}} \subset \mathcal{U} \times \mathcal{D}$, but it need not be a Cartesian product because the value of d affects that of f_1.

Lemma 5. *There exists a continuous function $\tilde{V} : X \times \tilde{\mathcal{Q}} \to \mathbb{R}$ such that, for all $f \in \mathcal{F}_0$,*

$$V(f) = \tilde{V}(f_0, f_1, d({}_2f)), \tag{9}$$

where d is the function defined in Lemma 2. For any f_1, f'_1, d', and d'' we have the following:
(5.i) if (f_1, d'') and (f'_1, d'') are in $\tilde{\mathcal{Q}}$, $\tilde{V}(f_0, f_1, d'') > \tilde{V}(f_0, f'_1, d'')$ iff[38] *$f_1 > f'_1$;*
(5.ii) if (f_1, d') and (f_1, d'') are in $\tilde{\mathcal{Q}}$, $\tilde{V}(f_0, f_1, d') > \tilde{V}(f_0, f_1, d'')$ iff $d' > d''$.

Proof. Recall that $d(\cdot)$ defines a ranking on $\mathcal{F}$ and that ${}_2f \in \mathcal{F}$. For any $(f_0, f_1, {}_2f)$ and $(f'_0, f'_1, {}_2f')$ such that both $(f_0, f_1, {}_2f')$ and $(f'_0, f'_1, {}_2f)$ are in $\mathcal{F}_0$, by Lemma 1 with $\pi = \{2, 3\}$, $V(f_0, f_1, {}_2f) \geq$

[38] Hereafter, "iff" stands for "if and only if."

$V(f_0, f_1, {}_2f')$ iff $V(f'_0, f'_1, {}_2f) \geq V(f'_0, f'_1, {}_2f')$. Moreover, for any $(f_1, {}_2f)$ and $(f_1, {}_2f')$ in $\mathcal{F}$, by Axiom 8, $V(f_0, f_1, {}_2f) \geq V(f_0, f_1, {}_2f')$ iff $W(u(f_0), d({}_2f)) \geq W(u(f_0), d({}_2f'))$, and therefore iff $d({}_2f) \geq d({}_2f')$. So, the ranking of $(f_0, f_1, {}_2f)$ and $(f_0, f_1, {}_2f')$ depends only on the value of $d(\cdot)$. Now, for any $f \in \mathcal{F}_0$, set

$$\tilde{V}(f_0, f_1, d({}_2f)) = V(f_0, f_1, {}_2f).$$

The previous argument implies property (5.i).

$\tilde{V}$ is well defined for the following reasons. First, if $(f_0, f_1, {}_2f)$ and $(f'_0, f'_1, {}_2f')$ are such that $f_t = f'_t$ for $t = 0, 1$ and $d({}_2f) = d({}_2f')$, then $V(f_0, f_1, {}_2f) = V(f'_0, f'_1, {}_2f')$ again by Axiom 8. Second, if $(f_0, f_1, {}_2f)$ and $(f'_0, f'_1, {}_2f')$ are such that either $(f_0, f_1, {}_2f') \notin \mathcal{F}_0$ or $(f'_0, f'_1, {}_2f) \notin \mathcal{F}_0$, then $(f_0, f_1) \neq (f'_0, f'_1)$. So, even if $d({}_2f) = d({}_2f')$, $\tilde{V}(f_0, f_1, d({}_2f))$ can be different from $\tilde{V}(f'_0, f'_1, d({}_2f'))$.

Consider now (f_1, d''), $(f'_1, d'') \in \tilde{\mathcal{Q}}$. There exist $c, c' \in C$, such that $f_t = f_t(c)$ and $f'_t = f_t(c')$ for $t = 0, 1$, and $d({}_2f(c)) = d({}_2f(c')) = d''$. By Lemma 2, without loss, we can assume that ${}_2c = {}_2c'$ so that ${}_2f(c) = {}_2f(c') = {}_2f''$. By Axiom 7(i), then $V(f_0, f_1, {}_2f'') > V(f_0, f'_1, {}_2f'')$ iff $f_1 > f'_1$, and property (5.ii) follows from (9).

Finally, $\tilde{V}$ is continuous for the following reasons. For any $(f_1, d) \in \tilde{\mathcal{Q}}$, $\tilde{V}(\cdot, f_1, d) = U(\cdot, {}_1c)$ for any c such that $f_1 = f_1(c)$ and $d = d({}_2f(c))$. Hence, the continuity property of U implies that $\tilde{V}$ is continuous in its first argument. Given any f_0 and $d \in \mathcal{D}$, $\tilde{V}(f_0, \cdot, d) = U(f_0, \cdot, {}_2c)$ for some c such that $f_0(c) = f_0$ and $d = d({}_2f(c))$. Hence, $\tilde{V}$ must take value in a connected interval and, being strictly increasing, it must be continuous in its second argument given f_0 and d. By a similar argument, for any (f_0, f_1), $\tilde{V}(f_0, f_1, \cdot)$ must take values in a connected interval and, being strictly increasing, it must be continuous in its last argument. It follows that $\tilde{V}$ must be continuous on the connected set $X \times \tilde{\mathcal{Q}}$.

□

Now define $\mathcal{Q} = \{(f_1(c), f_2(c), d({}_3f(c))) : c \in C\}$. By an argument similar to that in the proof of Lemma 5, using Lemma 1 with $\pi = \{3\}$, we obtain the following.

Lemma 6. *There exists a continuous function $\overline{V} : X \times \mathcal{Q} \to \mathbb{R}$ such that, for all $f \in \mathcal{F}_0$,*

$$V(f) = \overline{V}(f_0, f_1, f_2, d({}_3f)), \tag{10}$$

where d is the function defined in Lemma 2. Moreover, for any f_1, f'_1, f_2, f'_2, d', and d'' we have the following:
(6.i) if (f_1, f_2, d'), $(f'_1, f_2, d') \in \mathcal{Q}$, then $\overline{V}(f_0, f_1, f_2, d') > \overline{V}(f_0, f'_1, f_2, d')$ iff $f_1 > f'_1$;
(6.ii) if (f_1, f_2, d'), $(f_1, f'_2, d') \in \mathcal{Q}$, then $\overline{V}(f_0, f_1, f_2, d') > \overline{V}(f_0, f_1, f'_2, d')$ iff $f_2 > f'_2$;
(6.iii) if (f_1, f_2, d'), $(f_1, f_2, d'') \in \mathcal{Q}$, then $\overline{V}(f_0, f_1, f_2, d') > \overline{V}(f_0, f_1, f_2, d'')$ iff $d' > d''$.

Hereafter, for any $c \in C$, let $d_3(c) = d({}_3f(c))$. Also, we say that $c \in C$ induces $(f_0, f_1, f_2, d_3) \in X \times \mathcal{Q}$ if $f_t(c) = f_t$ for $t = 0, 1, 2$ and $d_3(c) = d_3$. Note that the function $\overline{V}$ defines a preference $\bar{\succ}$ on $X \times \mathcal{Q}$; moreover, by definition, for $c, c' \in C$

$$(f_0(c), f_1(c), f_2(c), d_3(c)) \bar{\succ} (f_0(c'), f_1(c'), f_2(c'), d_3(c')) \Leftrightarrow f(c) \succ^* f(c').$$

Lemma 7. *The preference $\bar{\succ}$ satisfies the following property (see Definition 4 in Debreu (1960)). Fix* *any nonempty subset* π *of* $\{0,1,2,3\}$*. Then*

$$(f_0, f_1, f_2, d_3)\succsim(f'_0, f'_1, f'_2, d'_3) \Leftrightarrow (\hat{f}_0, \hat{f}_1, \hat{f}_2, \hat{d}_3)\succsim(\hat{f}'_0, \hat{f}'_1, \hat{f}'_2, \hat{d}'_3)$$

where $f_t = \hat{f}_t$*,* $f'_t = \hat{f}'_t$*,* $d_3 = \hat{d}_3$*, and* $d'_3 = \hat{d}'_3$ *if* t *or* 3 *are in* π*, and* $f_t = f'_t$*,* $\hat{f}_t = \hat{f}'_t$*,* $d_3 = d'_3$*, and* $\hat{d}_3 = \hat{d}'_3$ *if* t *or* 3 *are not in* π*.*

Proof. Given π, let π^c be its complement. If $3 \in \pi^c$, then there exist $c, c', \hat{c}, \hat{c}' \in C$ such that, for $t = 0, 1, 2$, $f_t = f_t(c)$, $f'_t = f_t(c')$, $\hat{f}'_t = f_t(\hat{c}')$, $\hat{f}_t = f_t(\hat{c})$, $d({}_3f(c)) = d({}_3f(c'))$, and $d({}_3f(\hat{c})) = d({}_3f(\hat{c}'))$. Then, by Lemma 2, $f_2(c) = f_2(c_0, c_1, c_2, {}_3c')$ and $f_2(\hat{c}) = f_2(\hat{c}_0, \hat{c}_1, \hat{c}_2, {}_3\hat{c}')$. Similarly, by Lemma 5, $f_1(c) = f_1(c_0, c_1, c_2, {}_3c')$ and $f_1(\hat{c}) = f_1(\hat{c}_0, \hat{c}_1, \hat{c}_2, {}_3\hat{c}')$. Therefore, we can take ${}_3c = {}_3c'$ and ${}_3\hat{c} = {}_3\hat{c}'$, so that ${}_3f = {}_3f'$ and ${}_3\hat{f} = {}_3\hat{f}'$.[39] It follows from Lemma 1, that

$$V(f_0, f_1, f_2, {}_3f) > V(f'_0, f'_1, f'_2, {}_3f') \Leftrightarrow V(\hat{f}_0, \hat{f}_1, \hat{f}_2, {}_3\hat{f}) > V(\hat{f}'_0, \hat{f}'_1, \hat{f}'_2, {}_3\hat{f}').$$

Hence, by (10), the result follows.

Suppose $3 \in \pi$. Again, there exist $c, c', \hat{c}, \hat{c}' \in C$, each inducing the respective element of $X \times \mathcal{Q}$—in particular, $d({}_3f(c)) = d({}_3f(\hat{c}))$ and $d({}_3f(c')) = d({}_3f(\hat{c}'))$. Then, by Lemma 2, $f_2(c) = f_2(c_0, c_1, c_2, {}_3\hat{c})$ and $f_2(c') = f_2(c'_0, c'_1, c'_2, {}_3\hat{c}')$. Similarly, by Lemma 5, $f_1(c) = f_1(c_0, c_1, c_2, {}_3\hat{c})$ and $f_1(c') = f_1(c'_0, c'_1, c'_2, {}_3\hat{c}')$. Therefore, we can take ${}_3c = {}_3\hat{c}$ and ${}_3c' = {}_3\hat{c}'$, so that ${}_3f = {}_3\hat{f}$ and ${}_3f' = {}_3\hat{f}'$. It follows again from Lemma 1, that

$$V(f_0, f_1, f_2, {}_3f) > V(f'_0, f'_1, f'_2, {}_3f') \Leftrightarrow V(\hat{f}_0, \hat{f}_1, \hat{f}_2, {}_3\hat{f}) > V(\hat{f}'_0, \hat{f}'_1, \hat{f}'_2, {}_3\hat{f}').$$

Hence, by (10), the result follows.

□

Step 3: Let $\mathcal{O}$ be the set of vectors $(f_1(c), f_2(c), d({}_3f(c)))$ for $c \in C$, such that $\underline{u} < u(c_t) < \overline{u}$ for $t = 1, 2$ and $d({}_3f(c)) \in int\mathcal{D}$. It is straightforward to check that $\mathcal{O}$ is nonempty and that $\mathcal{Q}$ is included in the closure of $\mathcal{O}$.[40]

Lemma 8. *For any* $(f_1, f_2, d_3) \in \mathcal{O}$*, there exists* $\eta > 0$ *such that the rectangle*

$$\mathcal{R}(f_1, f_2, d_3; \eta) = (f_1 - \eta, f_1 + \eta) \times (f_2 - \eta, f_2 + \eta) \times (d_3 - \eta, d_3 + \eta)$$

lies in $\mathcal{O}$*.*

[39] Recall that by Lemma 6, if $(f_0, f_1, f_2, {}_3\tilde{f})$ and $(f_0, f_1, f_2, {}_3\tilde{f}')$ are in $\mathcal{F}_0$ and $d({}_3\tilde{f}) = d({}_3\tilde{f}')$, then $V(f_0, f_1, f_2, {}_3\tilde{f}) = V(f_0, f_1, f_2, {}_3\tilde{f}')$.

[40] To see that $\mathcal{O} \neq \varnothing$, consider any constant $c' \in C$ such that $\underline{u} < u(c'_0) < \overline{u}$. By changing c'_3 so that $u(c_3)$ varies continuously in an open interval around $u(c'_3)$, by continuity of U we can continuously span an open interval around $f_3(c')$. By Axiom 7(i), this variation in c_3 leads to variations in $f_2(c)$, which must span an open interval around $f_2(c')$, again by continuity of U. Since we are not changing c'_2, by Lemma 2, $d_3(c)$ must change in an open interval around $d_3(c')$. Finally, by Lemma 5, $f_1(c)$ must also vary continuously in an open interval around $f_1(c')$. To see that $\mathcal{Q} \subset cl\mathcal{O}$, notice that any point of $\mathcal{Q}$ induced by some $c \in C$ can be approximated, by slightly modifying c, by a c' such that $d_3(c') \in int\mathcal{D}$ and $u(c'_t) \in int\mathcal{I}_u$ for $t = 1, 2$, i.e., a point in $\mathcal{O}$.

Proof. Fix $(f_1, f_2, d_3) \in \mathcal{O}$ and, for the inducing c, let $u_t = u(c_t)$ for $t = 1, 2$. Since $d_3 \in int\mathcal{D}$, there is an interval $(\underline{d}_3, \overline{d}_3) \subset \mathcal{D}$ containing d_3. Since $\underline{u} < u_2 < \overline{u}$, given d_3, there is an interval $(\underline{f}_2(d_3), \overline{f}_2(d_3)) \subset \mathcal{U}$, containing f_2 and spanned by $u_2 \in int\mathcal{I}_u$. Let $\eta' > 0$ be such that $[d_3 - \eta', d_3 + \eta'] \subset (\underline{d}_3, \overline{d}_3)$. By the properties of W in Lemma 2, there exists $\eta' > 0$ such that $\underline{f}_2(d_3) < \underline{f}_2(d_3 + \eta') < f_2$ and $\overline{f}_2(d_3) > \overline{f}_2(d_3 - \eta') > f_2$. Hence, for all $d_3' \in [d_3 - \eta', d_3 + \eta']$, all $f_2' \in [f_2 - \varepsilon(\eta'), f_2 + \varepsilon(\eta')]$ are achievable by changing only u_2, where $\varepsilon(\eta') = \min\{f_2 - \underline{f}_2(d_3 + \eta'), \overline{f}_2(d_3 - \eta') - f_2\}$. Since $\underline{u} < u_1 < \overline{u}$, given f_2 and d_3, there is an interval $(\underline{f}_1(f_2, d_3), \overline{f}_1(f_2, d_3)) \subset \mathcal{U}$, containing f_1 and spanned by $u_1 \in int\mathcal{I}_u$. By the properties of $\tilde{V}$ in Lemma 5, there exist $\eta'' > 0$ and $\varepsilon'' > 0$ such that $[d_3 - \eta'', d_3 + \eta''] \subset (\underline{d}_3, \overline{d}_3)$, $[f_2 - \varepsilon'', f_2 + \varepsilon''] \subset (\underline{f}_2(d_3), \overline{f}_2(d_3))$, and $\underline{f}_1(f_2, d_3) < \underline{f}_1(f_2 + \varepsilon'', d_3 + \eta'') < f_1$ and $\overline{f}_1(f_2, d_3) > \overline{f}_1(f_2 - \varepsilon'', d_3 - \eta'') > f_1$. Hence, for all $(f_2'', d_3'') \in [f_2 - \varepsilon'', f_2 + \varepsilon''] \times [d_3 - \eta'', d_3 + \eta'']$, all $f_1'' \in [f_1 - \delta(\varepsilon'', \eta''), f_1 + \delta(\varepsilon'', \eta'')]$ are achievable by changing only u_1, where $\delta(\varepsilon'', \eta'') = \min\{f_1 - \underline{f}_1(f_2 + \varepsilon'', d_3 + \eta''), \overline{f}_1(f_2 - \varepsilon'', d_3 - \eta'') - f_1\}$. Let $\hat{\eta} = \min\{\eta', \eta''\}$, $\varepsilon = \min\{\varepsilon(\hat{\eta}), \varepsilon''\}$, and $\delta = \delta(\varepsilon, \hat{\eta})$. Noting that $\varepsilon(\hat{\eta}) \geq \varepsilon(\eta')$ and letting $\eta = \min\{\hat{\eta}, \varepsilon, \delta\}$, we have that all (f_1', f_2', d_3') in $\mathcal{R}(f_1, f_2, d_3; \eta)$ are induced by some $c \in C$ and belong to $\mathcal{O}$.

□

Lemma 9. *$\mathcal{O}$ is connected.*

Proof. We will show that $\mathcal{O}$ is path connected and hence connected. Take any $(f_1', f_2', d_3'), (f_1'', f_2'', d_3'') \in \mathcal{O}$ with inducing streams $c', c'' \in C$. By definition, $u(c_t'), u(c_t'') \in int\mathcal{I}_u$ for $t = 1, 2$ and $d_3', d_3'' \in int\mathcal{D}$. Since $\mathcal{D}$ is an interval, we can vary consumption from $t = 3$ onward, creating a path from ${}_3c'$ to ${}_3c''$ so as to cover the interval between d_3' and d_3''. Along this path d_3 remains in $int\mathcal{D}$; moreover, by Lemma 2, f_2 varies covering an interval between f_2' and $f_2(c_0', c_1', c_2', {}_3c'')$, and by Lemma 5, f_1 varies covering an interval between f_1' and $f_1(c_0', c_1', c_2', {}_3c'')$. Since c_1' and c_2' are unchanged, all (f_1, f_2, d_3) along the path are in $\mathcal{O}$. Now fix ${}_3c = {}_3c''$ and vary c_2 to create a path from c_2' to c_2'' so as to cover the interval between $u(c_2')$ and $u(c_2'')$. Along this path $u(c_2)$ remains in $int\mathcal{I}_u$; moreover, by Lemma 2, f_2 varies covering the interval between $f_2(c_0', c_1', c_2', {}_3c'')$ and $f_2(c_0', c_1', c_2'', {}_3c'')$, and by Lemma 5, f_1 varies covering an interval between $f_1(c_0', c_1', c_2', {}_3c'')$ and $f_1(c_0', c_1', c_2'', {}_3c'')$. Since c_1' is unchanged, again all (f_1, f_2, d_3) along this second path are in $\mathcal{O}$. Finally, fix ${}_2c = {}_2c''$ and vary c_1 to create a path from c_1' to c_1'' so as to cover the interval between $u(c_1')$ and $u(c_1'')$. Along this path $u(c_1)$ remains in $int\mathcal{I}_u$; moreover, by Lemma 2, f_1 varies covering the interval between $f_1(c_0', c_1', c_2'', {}_3c'')$ and $f_1(c_0', c_1'', c_2'', {}_3c'')$. Since c_2'' is unchanged, again all (f_1, f_2, d_3) along this third path are in $\mathcal{O}$. The three paths together form a connected path from (f_1', f_2', d_3') to (f_1'', f_2'', d_3'') which never leaves $\mathcal{O}$.

□

We are now ready to obtain an additive representation of $\bar{\succ}$, relying on Debreu (1960).

Lemma 10. *The preference $\bar{\succ}$ over $X \times \mathcal{Q}$ can be represented by an additive function*

$$V^0(f_0, f_1, f_2, d_3) = \hat{u}(f_0) + a(f_1) + b(f_2) + \zeta(d_3),$$

where $\hat{u}$, a, b, and ζ are continuous, and a, b, ζ are strictly increasing on $\mathcal{U}$.

Proof. We first show that $\bar{\succ}$ has an additive representation over $X \times \mathcal{O}$. By continuity, we then extend this representation to $X \times \mathcal{Q}$. The representation of $\bar{\succsim}$ over $X \times \mathcal{Q}$ immediately implies the desired representation of $\succ^*$ on $\mathcal{F}_0$.

The set $\mathcal{O}$ may be expressed as a countable union of open rectangles $\{\mathcal{R}^i\}_{i\in\mathbb{N}}$ of the form in Lemma 8, and such that for any j there is an $i < j$ such that $\mathcal{R}^i \cap \mathcal{R}^j \neq \varnothing$. To construct $\{\mathcal{R}^i\}_{i\in\mathbb{N}}$, proceed as follows. Let $\{\overline{\mathcal{B}}^n\}_{n=1}^{\infty}$ be the sequence of closed balls of radius n centered at the origin in $\mathbb{R}^3$. Then, let

$$\mathcal{K}^n = \{o \in \mathcal{O} \mid o \in \overline{\mathcal{B}}^n, \mathcal{B}^{1/n}(o) \subset \mathcal{O}\},$$

where $\mathcal{B}^{1/n}(o)$ is the open ball of radius $1/n$ centered at point o. For each n, $\mathcal{K}^n$ is compact[41] and the increasing sequence $\{\mathcal{K}^n\}_{n=1}^{\infty}$ converges to $\mathcal{O}$. So, each $\mathcal{K}^n$ can be covered by finitely many rectangles of the form in Lemma 8. Since $\mathcal{K}^n \subset \mathcal{K}^{n+1}$, when moving from $\mathcal{K}^n$ to $\mathcal{K}^{n+1}$, one can cover $\mathcal{K}^{n+1}$ by simply adding rectangles to those used to cover $\mathcal{K}^n$. Without loss, any added rectangle contains a point with rational coordinates not contained in other rectangles, so that the list of rectangles needed to cover $\mathcal{O}$, denoted by $\{\mathcal{R}^i\}_{i\in\mathbb{N}}$, is countable. Finally, since $\mathcal{O}$ is connected, each $\mathcal{R}^j$ must intersect at least another $\mathcal{R}^i$. For simplicity, we can relabel the rectangles so that, for each j, we have $\mathcal{R}^j \cap \mathcal{R}^i \neq \varnothing$ for some $i < j$.

For any $\mathcal{R}^i$, Lemmas 4 and 7 guarantee that the hypotheses of Debreu's (1960) Theorem 3 are satisfied on $X \times \mathcal{R}^i$. Therefore, $\bar{\succsim}$ may be expressed over each $X \times \mathcal{R}^i$ as

$$V^i(f_0, f_1, f_2, d_3) = \hat{u}^i(f_0) + a^i(f_1) + b^i(f_2) + \zeta^i(d_3),$$

for functions $\hat{u}^i, a^i, b^i$, and ζ^i that are continuous and, except for $\hat{u}^i$, strictly increasing by the properties of $\bar{V}$ which induces $\bar{\succsim}$.[42]

By construction, $\mathcal{R}^0$ and $\mathcal{R}^1$ have a nonempty open intersection. Over $X \times (\mathcal{R}^0 \cap \mathcal{R}^1)$ representations V^0 and V^1 must be positive affine transformations of each other (Debreu's (1960) Theorem 3). So there exist constants $\rho > 0$ and $\chi \in \mathbb{R}$ such that, on $X \times (\mathcal{R}^0 \cap \mathcal{R}^1)$,

$$\hat{u}^0(f_0) = \rho\hat{u}^1(f_0) + \chi, \quad a^0(f_1) = \rho a^1(f_1), \qquad\qquad b^0(f_2) = \rho b^1(f_2), \quad \zeta^0(d_3) = \rho\zeta^1(d_3).$$

Using these conditions, we can extend $\hat{u}^0, a^0, b^0$, and ζ^0 to the set $X \times (\mathcal{R}^0 \cup \mathcal{R}^1)$. Indeed, each function a^i, b^i, and ζ^i is defined on $\mathcal{R}^i_k$ which denotes the projection of $\mathcal{R}^i$ on the k^{th} dimension of $\mathcal{Q}$. Consider a^0. By extending a^0 over $\mathcal{R}^1_1 \setminus \mathcal{R}^0_1$ using a^1, the resulting function a^0 is well defined and continuous on $\mathcal{R}^0_1 \cup \mathcal{R}^1_1$. By a similar reasoning for b^0 and ζ^0, the function V^0 can be extended to $X \times (\mathcal{R}^0_1 \cup \mathcal{R}^1_1) \times (\mathcal{R}^0_2 \cup \mathcal{R}^1_2) \times (\mathcal{R}^0_3 \cup \mathcal{R}^1_3)$. Since this product includes $X \times \mathcal{R}^0 \cup \mathcal{R}^1$, the function V^0 is, in particular, well defined and continuous on it.

Finally, since for each $j > 0$ we have $\mathcal{R}^j \cap \mathcal{R}^i \neq \varnothing$ for some $i < j$, we can extend by induction representation V^0 from $X \times \mathcal{R}^0$ to $X \times (\cup_{i\in\mathbb{N}}\mathcal{R}^i) = X \times \mathcal{O}$, in countably many steps. Notice that the functions a, b, and ζ (we henceforth omit the superscript '0') entering the formula of V^0 are defined, through the induction, over the respective projections of $\mathcal{O}$.

[41] $\mathcal{K}^n$ is clearly bounded. Consider any sequence $\{o^m\} \subset \mathcal{K}^n$ converging to o'. Since $\overline{\mathcal{B}}^n$ is closed, $o' \in \overline{\mathcal{B}}^n$. There remains to show that $\mathcal{B}^{1/n}(o') \subset \mathcal{O}$. Let o'' be any point such that $||o' - o''|| = r < 1/n$. Then $||o'' - o^m|| \leq r + ||o' - o^m||$. So, for m large enough, $o'' \in \mathcal{B}^{1/n}(o^m)$ and hence $o'' \in \mathcal{O}$.

[42] While Debreu's theorem requires that the preference domain be a Cartesian product, it does not assume compactness of the sets forming the product.

Since any point of $X \times \mathcal{O}$ is contained in $X \times \mathcal{R}^i$ for some $i \in \mathbb{N}$, V^0 and its components $\hat{u}$, a, b, and ζ are continuous over $X \times \mathcal{O}$. Moreover, V^0 represents $\bar{\succ}$ on $X \times \mathcal{O}$. To see this, we need to check that for any (f_0', f_1', f_2', d_3') and $(f_0'', f_1'', f_2'', d_3'')$ in $X \times \mathcal{O}$, $V^0(f_0', f_1', f_2', d_3') > V^0(f_0'', f_1'', f_2'', d_3'')$ iff $(f_0', f_1', f_2', d_3') \bar{\succ} (f_0'', f_1'', f_2'', d_3'')$. Note that (f_1', f_2', d_3') and (f_1'', f_2'', d_3'') must both belong to some $\mathcal{K}^n$ in the previous construction. Since V^0 represents $\bar{\succ}$ on $X \times \mathcal{K}^n$, it ranks (f_0', f_1', f_2', d_3') and $(f_0'', f_1'', f_2'', d_3'')$ correctly, which proves the claim.

It remains to show that V^0 can be extended to the entire domain $X \times \mathcal{Q}$, additively, and that it represents $\bar{\succ}$ over this domain. We first show that V^0 can be extended to a continuous function over $X \times \mathcal{Q}$. Recall that $\overline{V}$ is continuous and represents $\bar{\succ}$ over $X \times \mathcal{Q}$—and hence over $X \times \mathcal{O}$. So, there exists a strictly increasing map $\phi : Y \to Y^0$ such that $V^0 = \phi \circ \overline{V}$, where Y^0 and Y are the ranges of V^0 and $\overline{V}$ on $X \times \mathcal{O}$. Y^0 and Y are intervals of $\mathbb{R}$ because $X \times \mathcal{O}$ is connected and V^0 and $\overline{V}$ are continuous over this domain. Since ϕ is strictly increasing, it must be continuous on its domain, otherwise it would not cover Y^0. Let $\overline{Y}$ be the range of $\overline{V}$ over $X \times \mathcal{Q}$. Since $X \times \mathcal{Q} \subset cl(X \times \mathcal{O})$ and $\overline{V}$ is continuous, $\overline{Y}$ contains at most two more points than Y (its boundaries), and this may occur only when the relevant boundaries are finite. One can extend ϕ to these points, whenever applicable, by taking the limit of ϕ: for example, if $\bar{y}$ denotes the upper bound of $\overline{Y}$ and $\bar{y} \notin Y$, one may define $\phi(\bar{y})$ as $\lim_{y \uparrow \bar{y}} \phi(y)$.[43] Finally, we can extend V^0 to $X \times \mathcal{Q}$ by letting $V^0 = \phi \circ \overline{V}$ over this domain. By construction, V^0 is continuous as the composition of continuous functions.

Next, we show that this extension of V^0 to $X \times \mathcal{Q}$ still obeys the additive representation obtained on $X \times \mathcal{O}$ in terms of $\hat{u}, a, b$ and ζ. We first show that a, b, and ζ can be extended on the relevant projections of $\mathcal{Q}$ (not just of $\mathcal{O}$). Since $\mathcal{O}$ is connected and $\mathcal{Q} \subset cl\mathcal{O}$, the extension is only needed (possibly) at the two boundaries of $\mathcal{D}$ for ζ, and at the boundaries of $\mathcal{U}$ for a and b; these extensions are necessary only if these boundaries are achieved by some $(f_1, f_2, d_3) \in \mathcal{Q}$.

To extend ζ, suppose that there is an $(f_1, f_2, d_3) \in \mathcal{Q}$ such that d_3 is the upper bound of $\mathcal{D}$—the other case follows similarly. Without loss, we can choose f_1, $f_2 \in int\mathcal{U}$.[44] By perturbing c_3, we can then construct a sequence $\{(f_1^n, f_2^n, d_3^n)\}$ such that f_1^n and f_2^n are in some compact $K \subset int\mathcal{U}$ and $d_3^n \in int\mathcal{D}$ for all n, and $d_3^n \to d_3$. By construction, each $(f_1^n, f_2^n, d_3^n) \in \mathcal{O}$. Fixing some f_0, the sum $\hat{u}(f_0) + a(f_1^n) + b(f_2^n) + \zeta(d_3^n)$ is well defined and equal to $V^0(f_0, f_1^n, f_2^n, d_3^n)$ for each n. Moreover, possibly moving to subsequences, $f_1^n \to \hat{f}_1$ and $f_2^n \to \hat{f}_2$ for some $\hat{f}_1, \hat{f}_2 \in K$. Since a and b are

[43] One can show that for $\bar{y} \in \overline{Y} \setminus Y$, $\lim_{y \uparrow \bar{y}} \phi(y)$ must be finite. Suppose not: First, there exist i) $\bar{s} = (\bar{f}_0, \bar{f}_1, \bar{f}_2, \bar{d}) \in X \times \mathcal{Q}$ such that $\bar{V}(\bar{s}) = \bar{y}$, which means that the agent prefers $\bar{s}$ to any other stream; and ii) a sequence $s^n = (f_0^n, f_1^n, f_2^n, d^n) \in X \times \mathcal{O}$ that converges to $\bar{s}$, and such that $V^0(s^n)$ diverges to $+\infty$. Because V^0 is additive, this means that there must be at least one sequence, among $\hat{u}(f_0^n)$, $a(f_1^n)$, $b(f_2^n)$, and $\zeta(d^n)$, with a subsequence diverging to $+\infty$. For example, suppose that d^n is such that $\zeta(d^n)$ diverges to $+\infty$. Then, for any stream c such that $d({}_3f(c)) = \bar{d}$, we have $\bar{V}(f_0(c), f_1(c), f_2(c), \bar{d}(c)) = \bar{y}$. Indeed, fix any c_0 and c_t's such that $\underline{u} < u(c_t) < \bar{u}$ for $t \in 1, 2$. Choosing the sequence of continuation streams $(c_3, \ldots)$ corresponding to the sequence of d^n converging to $\bar{d}$, V^0 evaluated at those streams (and the fixed c_0, c_1, c_2) must diverge to $+\infty$. This implies that $\bar{V}$ converges to $\bar{y}$ for that sequence. By continuity of $\bar{V}$ over its entire domain, this implies that when choosing $(c_3, \ldots)$ such that $d({}_3f(c)) = \bar{d}$, we have $\bar{V}(f_0(c), f_1(c), f_2(c), d({}_3f(c))) = \bar{y}$, regardless of the values of c_0, c_1, c_2. This, however, violates the fact that preferences are strictly increasing in $u(c_0)$ (Lemma 2), a contradiction. A similar contradiction can be derived if instead $\hat{u}(f_0^n)$, or $a(f_1^n)$, or $b(f_2^n)$ has a subsequence diverging to $+\infty$. This shows that necessarily Y_0 is bounded above whenever $\bar{y} \in \bar{Y} \setminus Y$. By a similar argument for the lower bound, we conclude that ϕ is bounded at any boundary for which it needs to be extended.

[44] This can be achieved by changing c_1 and c_2 of the stream inducing (f_1, f_2, d_3), without affecting d_3.

continuous over K, $a(f_1^n)$ and $b(f_2^n)$ converge on these subsequences. Therefore, $\zeta(d_3)$ is well defined as the difference $V^0(f_0, \hat{f}_1, \hat{f}_2, d_3) - \hat{u}(f_0) - a(\hat{f}_1) - b(\hat{f}_2)$, because V^0 has already been extended to $(f_0, \hat{f}_1, \hat{f}_2, d_3)$. Moreover, since V^0 was extended continuously over $X \times \mathcal{Q}$, ζ must also be continuous at d_3.

We can similarly extend b to the boundary of $\mathcal{U}$, whenever needed. To see this, take any $(f_1, f_2, d_3) \in \mathcal{Q}$ such that f_2 lies at a boundary of $\mathcal{U}$, say $f_2 = \bar{\nu}$—again, the other case follows similarly. Moreover, we can choose c_1 in the inducing stream $c \in C$ so that $f_1 \in int\mathcal{U}$. By perturbing c_2, we can build a sequence $\{(f_1^n, f_2^n, d_3)\}$ such that f_1^n is in a compact $K \subset int\mathcal{U}$ and $f_2^n \in int\mathcal{U}$ for all n, and $f_2^n \to \bar{\nu}$. Possibly taking a subsequence such that $f_1^n \to \hat{f}_1$ for some $\hat{f}_1 \in K$, we obtain a well define limit for V^0, a, and ζ, from which we can obtain the value of $b(\bar{\nu})$. The argument for a is identical.

In conclusion, the function $\hat{u}(\cdot) + a(\cdot) + b(\cdot) + \zeta(\cdot)$ is equal to V^0 over the entire set $X \times \mathcal{Q}$, and represents $\bar{\succ}$ over this domain.

$\square$

Step 4: By Lemma 1 with $\pi = \{1, 2, 3\}$, for any $f_0 \in X$, the induced preference $\succ_{-0}^*$ on $\mathcal{F}$ is independent of f_0. By Lemma 10, we can conclude that $\succ_{-0}^*$ has a representation

$$V_{-0}^*(f_1, f_2, {}_3f) = a(f_1) + b(f_2) + \zeta(d({}_3f)). \tag{11}$$

Note that Axiom 8 holds for any f_0. So if $f_1 = f_1'$, $(f_1, f_2, {}_3f) \succsim_{-0}^* (f_1', f_2', {}_3f')$ iff $(f_2, f_3, {}_4f) \succsim_{-0}^* (f_2', f_3', {}_4f')$.

Lemma 11. *There exist $\alpha > 0$, $\xi \in \mathbb{R}$, and $G : \mathcal{U} \to \mathbb{R}$ continuous and strictly increasing such that, for any finite $T \geq 2$ any $f \in \mathcal{F}$,*

$$V_{-0}^*(f) = \sum_{t=1}^{T} \alpha^t G(f_t) + \alpha^T \tilde{d}({}_{T+1}f) + \xi \sum_{t=0}^{T-2} \alpha^t. \tag{12}$$

Proof. Consider again $\mathcal{R}^0$ in the proof of Lemma 10. By definition of a rectangle, if $(f_1, f_2, {}_3f)$ and $(f_1', f_2', {}_3f')$ are such that $(f_1, f_2, d({}_3f))$, $(f_1', f_2', d({}_3f')) \in \mathcal{R}^0$, then all $\hat{f}_1 \in \mathcal{R}_1^0$ are feasible with both $(f_2, {}_3f)$ and $(f_2', {}_3f')$. By the stationarity property of $\succ_{-0}^*$, we have

$$a(\hat{f}_1) + b(f_2) + \zeta(d({}_3f)) \geq a(\hat{f}_1 Let) + b(f_2') + \zeta(d({}_3f'))$$

if and only if

$$a(f_2) + b(f_3) + \zeta(d({}_4f)) \geq a(f_2') + b(f_3') + \zeta(d({}_4f')).$$

Hence, since additive representations are unique up to positive affine transformations, for all $(f_2, {}_3f)$ such that $(f_1, f_2, d({}_3f)) \in \mathcal{R}^0$,

$$\alpha(a(f_2) + b(f_3) + \zeta(d({}_4f))) + \xi = b(f_2) + \zeta(d({}_3f)) \tag{13}$$

for some $\alpha > 0$ and $\xi \in \mathbb{R}$.

The argument used for $\mathcal{R}^0$ can be equivalently applied to any $\mathcal{R}^i$ in the covering $\{\mathcal{R}^i\}_{i \in \mathbb{N}}$ of $\mathcal{O}$. Moreover, since for each $j > 0$ we have $\mathcal{R}^j \cap \mathcal{R}^i \neq \varnothing$ for some $i < j$, it is clear that the α in (13) must be the same for all $f \in \mathcal{F}$ such that $(f_1, f_2, d({}_3f)) \in \mathcal{O}$. That (13) must hold for all $f \in \mathcal{F}$ is implied by the following two observations. First, if $c \in C$ induces $(f_1, f_2, d_3) \in \mathcal{O}$, it imposes no restriction on $d({}_4f(c))$, which can take any value in $\mathcal{D}$—hence ${}_4f$ can take any value in $\mathcal{F}$. To see this, recall that for $f \in \mathcal{F}$ we defined $d(f) = V(\hat{f}_0, f)$ for some $\hat{f}_0 \in X$, and $V(\hat{f}_0, f) = \tilde{V}(\hat{f}_0, f_1, d({}_2f))$ by Lemma 5. So, since $\tilde{V}$ is strictly increasing in its second and third argument, the condition $d_3(c) \in int\mathcal{D}$ only implies $f_3(c) \in int\mathcal{U}$, but $d({}_4f(c))$ can be at the boundary of $\mathcal{D}$. Therefore, (13) already holds for any value of ${}_4f \in \mathcal{F}$. Second, suppose that f is such that $(f_1, f_2, d({}_3f))$ is at boundary of $\mathcal{Q}$. Take a sequence $\{f^n\}$ such that $(f_1^n, f_2^n, d({}_3f^n)) \in \mathcal{O}$ for all n and converges to $(f_1, f_2, d({}_3f))$. The sequence can be chosen so that ${}_4f$ is fixed: perturbing only c_1, c_2, and c_3 is enough to guarantee that we are in $\mathcal{O}$. Now recall that the functions a, b, and ζ are continuous by Lemma 10. Then, the right-hand side of (13) converges, as do the first two terms of the left-hand side. The last term is constant and equal to $\zeta(d({}_4f))$, so it converges trivially. Therefore (13) holds everywhere.

We conclude that, for all $f \in \mathcal{F}$,

$$V^*_{-0}(f_1, f_2, {}_3f) = a(f_1) + \xi + \alpha V^*_{-0}(f_2, f_3, {}_4f).$$

Therefore, using this condition recursively and (11), for any $f \in \mathcal{F}$ and finite $T > 2$, we have

$$V^*_{-0}(f_1, f_2, {}_3f) = \sum_{t=0}^{T-1} \alpha^t a(f_{t+1}) + \alpha^{T-1}(b(f_{T+1}) + \zeta(d({}_{T+2}f))) + \xi \sum_{t=0}^{T-2} \alpha^t.$$

The result then follows by defining $G = \alpha^{-1}a$ and $\tilde{d}(\cdot) = \alpha^{-1}(b(\cdot) + \zeta(d(\cdot)))$.

□

By Lemma 11, for any finite $T \geq 2$, we can represent $\succ$ for streams c as

$$\overline{U}(c) = \hat{u}(c_0) + \sum_{t=1}^{T} \alpha^t G(U({}_tc)) + \alpha^T \tilde{d}(U({}_{T+1}c), U({}_{T+2}c), \ldots) + \xi \sum_{t=0}^{T-2} \alpha^t. \tag{14}$$

The next two technical lemmas will be useful to complete the proof of our theorem.

Lemma 12. *For any constant streams $c, c' \in C$, $c \succ c'$ iff $\hat{u}(c_0) > \hat{u}(c'_0)$.*

Proof. Suppose $\hat{u}(x) > \hat{u}(y)$ and consider $c = (x, x, \ldots)$ and $\hat{c} = (x, y, \ldots)$. For any $t \geq 0$ and $c'' \in C$, define $c^t = (c_0, \ldots, c_t, c'')$ and $\hat{c}^t = (\hat{c}_0, \ldots, \hat{c}_t, c'')$. For $t = 0$, we have $\overline{U}(c^t) = \overline{U}(\hat{c}^t)$. For any $t \geq 1$, using (14), we first have $\overline{U}({}_tc^t) > \overline{U}({}_t\hat{c}^t)$. Then, using again (14) backward recursively and monotonicity of G, we conclude that $\overline{U}(c^t) \geq \overline{U}(\hat{c}^t)$. Since this is true for any $t \geq 0$ and $c'' \in C$, Axiom 7(ii) implies $c \succsim \hat{c}$. Now note that, again by (14), $\hat{c} \succ (y, y, \ldots)$. Hence, by Axiom 1, $c \succ (y, y, \ldots)$.

Now suppose $\hat{u}(x) = \hat{u}(y)$ and consider $c = (x, x, \ldots)$ and $\hat{c} = (y, y, \ldots)$. For any t and $c'' \in C$, define c^t and $\hat{c}^t$ as before. Using again (14) backward recursively and the fact that G is a function, we conclude that $\overline{U}(c^t) = \overline{U}(\hat{c}^t)$. Since this is true for any t and $c'' \in C$, Axiom 7(ii) implies $c \sim c'$.

□

Lemma 13. *For any $c \in C$, there exists $x \in X$ such that $c \sim (x, x, \ldots)$.*

Proof. By Lemma 19 in the Supplemental Material, for any $c \in C$, there exists $y \in X$ such that $c \sim (c_0, y, y, \ldots)$. Suppose $(c_0, y, y, \ldots) \not\sim (y, y, \ldots)$. If $(c_0, y, y, \ldots) \succ (y, y, \ldots)$, then $\hat{u}(c_0) > \hat{u}(y)$. Let $\hat{c} = (c_0, c_0, \ldots)$ and $\tilde{c} = (c_0, y, y, \ldots)$. For any $t \geq 0$ and any $c'' \in C$, consider $\hat{c}^t = (\hat{c}_0, \ldots, \hat{c}_t, c'')$ and $\tilde{c}^t = (\tilde{c}_0, \ldots, \tilde{c}_t, c'')$. We have $\hat{c}^t \succsim \tilde{c}^t$. Indeed, for $t = 0$, $\hat{c}^t = \tilde{c}^t$. For $t \geq 1$, we can proceed using (14). Since $\hat{u}(c_0) > \hat{u}(y)$, $U({}_t\hat{c}^t) > U({}_t\tilde{c}^t)$. For $s < t$, since $\hat{u}(\hat{c}^t_s) \geq \hat{u}(\tilde{c}^t_s)$ and G is strictly increasing, we have $U({}_s\hat{c}^t) \geq U({}_s\tilde{c}^t)$. By Axiom 7(ii), we then have $\hat{c} \succsim \tilde{c}$ and hence $(c_0, c_0, \ldots) \succsim c \succ (y, y, \ldots)$. Since X is connected, by Axiom 2, there exists $x \in X$ such that $(x, x, \ldots) \sim c$. The case $(c_0, y, y, \ldots) \prec (y, y, \ldots)$ follows similarly.

□

We can now prove that $\alpha < 1$.

Lemma 14. $\alpha < 1$.

Proof. Consider consumption streams that are constant from $t = 3$ onward. Then f_t is constant for $t \geq 3$. So we can write $d({}_3f) = d({}_4f) = e(f_3)$ in (13) and thus obtain

$$(1 - \alpha)e(f_3) = \alpha b(f_3) + \alpha a(f_2) - b(f_2) + \xi.$$

First, note that $f_3 > f'_3$ implies $e(f_3) > e(f'_3)$. By Lemma 12, $f_3 > f'_3$ implies $u(c_3) > u(c'_3)$. Define $c = (c_3, c_3, \ldots)$ and $c' = (c_3, c'_3, c'_3, \ldots)$. Replicating the argument in the proof of Lemma 13, we have $c \succsim (c_3, c_3, c'_3, c'_3, \ldots)$. Moreover, by Axiom 7(i), $(c_3, c_3, c'_3, c'_3, \ldots) \succ c'$. Then, by Axiom 1 and Lemma 2, $W(u(c_3), d(f_3, f_3, \ldots)) > W(u(c_3), d(f'_3, f'_3, \ldots))$, which holds iff $d(f_3, f_3, \ldots) > d(f'_3, f'_3, \ldots)$.

Second, we can find $\hat{c}, \tilde{c} \in C$, constant from $t = 3$ onward, such that $f_2(\hat{c}) = f_2(\tilde{c})$ and $f_3(\hat{c}) > f_3(\tilde{c})$. Consider $x, y \in X$ with $u(x) > u(y)$ and the streams $(x, y, y, \ldots)$ and $(y, x, x, \ldots)$. By the previous argument based on Axiom 7(ii), $(x, x, x, \ldots) \succ (x, y, y, \ldots)$. If $(x, y, y, \ldots) \succsim (y, x, x, \ldots)$, then by (14) and continuity of $\hat{u}$ there exists $z \in X$ such that $(x, y, y, \ldots) \sim (z, x, x, \ldots)$. In this case, let $\hat{c} = (c_0, c_1, z, x, x, \ldots)$. If $(x, y, y, \ldots) \prec (y, x, x, \ldots)$, then by Axiom 2 there exists $w \in X$ such that $(x, y, y, \ldots) \sim (y, w, w, \ldots)$. Moreover, $u(w) > u(y)$. Otherwise, since $(y, y, y, \ldots) \succsim (y, w, w, \ldots)$ for $u(y) \geq u(w)$ (again by the same argument as before), we would have $(x, y, y, \ldots) \succ (y, w, w, \ldots)$ by (14) and Axiom 1. In this case, let $\hat{c} = (c_0, c_1, y, w, w, \ldots)$. Finally, let $\tilde{c} = (c_0, c_1, x, y, y, \ldots)$.

To conclude the proof, note that for $c \in \{\hat{c}, \tilde{c}\}$, $(1 - \alpha)e(f_3(c)) = \alpha b(f_3(c)) + \xi'$ for some constant ξ'. Since b and e are strictly increasing, we must have $\alpha < 1$.

□

Note that, if c is constant from any $T \geq 3$ onward, by Lemma 14 and (13)

$$\tilde{d}(U({}_Tc), U({}_{T+1}c), \ldots) = \frac{\alpha}{1 - \alpha} G(U({}_Tc)) + \frac{\xi}{\alpha(1 - \alpha)}.$$

So, for eventually constant streams, we can write

$$\overline{U}(c) = \hat{u}(c_0) + \sum_{t=1}^{T} \alpha^t G(U({}_tc)) + \frac{\alpha^{T+1}}{1 - \alpha} G(U({}_{T+1}c)) + \frac{1 + \alpha(1 - \alpha^{T-1})}{\alpha(1 - \alpha)} \xi. \tag{15}$$

Lemma 15. *G is bounded on $\mathcal{U}$.*

Proof. By Axiom 3, V^*_{-0} is finite for all $c \in C$. Suppose that G is unbounded above—the other case follows similarly. Then, for each $r \in \mathbb{R}_{++}$, there must be a stream c^r with utility U^r such that $G^r \equiv G(U^r) \geq r$. Moreover, for $r > r'$, we can choose c^r and $c^{r'}$ so that $G^r > G^{r'}$, relying on continuity of G and connectedness of $\mathcal{U}$. By Lemma 13, for each r we can also let c^r be constant. As a preliminary observation, note the following: given $r' > r$, a stream c that equals c^r for the first k periods and $c^{r'}$ forever after must satisfy $G(U(c)) \geq r$. This is because, by definition, $U({}_tc) > U({}_tc^r)$ for $t \geq k$; then, by monotonicity of G and using (15) backward recursively, we have $U({}_tc) > U({}_tc^r)$ for $0 \leq t < k$.

Now construct stream $\hat{c}$ as follows. For some $M > 1$ and each $t \geq 1$, consider the constant stream $c^{(M/\alpha)^t}$ with the property $\alpha^t G^{(M/\alpha)^t} \geq M^t$. Then, let $\hat{c}_0$ be such that $\underline{u} < \hat{u}(\hat{c}_0) < \overline{u}$ and, for each $t \geq 2$, let $\hat{c}_t = c_t^{(M/\alpha)^t}$. Now, for any $T > 0$, let c^T be equal to $\hat{c}$ up to T and to $c^{(M/\alpha)^T}$ thereafter. Using (15), we have

$$\begin{aligned}\overline{U}(c^T) =& \hat{u}(\hat{c}_0) + \sum_{t=1}^{T-1} \alpha^t G(U({}_tc^T)) + \frac{\alpha^T}{1-\alpha} G^{(M/\alpha)^T} + \frac{1+\alpha(1-\alpha^{T-2})}{\alpha(1-\alpha)}\xi \\ \geq& \hat{u}(\hat{c}_0) + \sum_{t=1}^{T-1} M^t + \frac{1}{1-\alpha} M^T + \frac{1+\alpha(1-\alpha^{T-2})}{\alpha(1-\alpha)}\xi,\end{aligned}$$

where the inequality follows by recursively applying our preliminary observation. Note that the lower bound on $\overline{U}(c^T)$ goes to $+\infty$ as $T \to \infty$.

Now fix any T and c^T. To simplify notation, let $\tilde{c} = c^T$. Using Axiom 7(ii), we have $\overline{U}(\hat{c}) \geq \overline{U}(\tilde{c})$. To see this, consider any $t \geq 0$ and $c'' \in C$, and let $\hat{c}^t = (\hat{c}_0, \ldots, \hat{c}_t, c'')$ and $\tilde{c}^t = (\tilde{c}_0, \ldots, \tilde{c}_t, c'')$. For $t \leq T$, we have $\hat{c}^t \sim \tilde{c}^t$ because the two streams are identical. For $t \geq T+1$, we first have that $u(\hat{c}_s) > u(\tilde{c}_s)$ for $T < s \leq t$ by Lemma 12. Hence, $\overline{U}(\hat{c}_t, c'') > \overline{U}(\tilde{c}_t, c'')$. Second, using again monotonicity of G and (12) recursively, we conclude $\overline{U}(\hat{c}^t) \geq \overline{U}(\tilde{c}^t)$. By Axiom 7(ii), we then have the claimed property.

It follows that, for any T, $\overline{U}(\hat{c}) \geq \overline{U}(c^T)$ and hence, since $\hat{u}(\hat{c}_0)$ is bounded by assumption, $V^*_{-0}(f(\hat{c}))$ must be infinite, violating Axiom 3.

□

Lemma 16. *For any $c \in C$, $\overline{U}(c) = \hat{u}(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U({}_tc))$.*

Proof. Again by Axiom 3, V^*_{-0} is finite for all $c \in C$. Using (12) for any finite T and observing that ${}_Tf$ can take any value in $\mathcal{F}$, we conclude that the function $\tilde{d}$ must be finite because G is bounded. The result then follows by letting $T \to \infty$, relying on $\alpha < 1$ and ignoring the additive constant.

□

To conclude, both functions U and $\overline{U}$ represent $\succ$ over C. So, they are strictly increasing transformations of one another. Letting $\overline{G}$ denote the function of $\overline{U}$ such that $\overline{G}(\overline{U}(c)) = G(U(c))$ for all c, we obtain representation (5). For uniqueness, note that the additive form of $\overline{U}$ is unique up to positive affine transformations, i.e., $\tilde{U} = \rho\overline{U} + \chi$ for $\rho > 0$ and $\chi \in \mathbb{R}$. So,

$$\tilde{U}(c) = \rho\hat{u}(c_0) + \chi + \sum_{t=1}^{\infty} \alpha^t \rho G(\overline{U}({}_tc)) = \rho\hat{u}(c_0) + \chi + \sum_{t=1}^{\infty} \alpha^t \rho G\left(\frac{\tilde{U}({}_tc) - \chi}{\rho}\right).$$

A.3 Proof of Proposition 1

Part (i). Take $\nu', \nu \in \mathcal{U}$. By definition, there exist $c', c \in C$ such that $U(c') = \nu'$ and $U(c) = \nu$. By Lemma 13, we can take $c' = (x, x, \dots)$ and $c = (y, y, \dots)$ for some $x, y \in X$. Suppose $u(x) > u(y)$. Then, by Lemma 12, $U(x, \dots) > U(y, \dots)$. 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 $\mathcal{U}$

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

Lemma 17. *For any $\varepsilon > 0$, there exists a constant $K \in (\frac{1-\alpha}{2\alpha}, \frac{1-\alpha}{\alpha})$ such that, for all $\nu' > \nu$ in $\mathcal{U}$,*

$$G(\nu') - G(\nu) \leq \max\{K(\nu' - \nu), \varepsilon\}$$

Proof. See the Supplemental Material.

□

To show that U is H-continuous, consider any $c, \tilde{c} \in C$ and define $c^T = (c_0, c_1, \dots, c_T, c)$ and $\tilde{c}^T = (c_0, c_1, \dots, c_T, \tilde{c})$. Using Lemma 17, we will show that for any $\varepsilon > 0$, there exists T such that

$$|U(cT) - U(\tilde{c}T)| < \frac{2\alpha\varepsilon}{1-\alpha}.$$

To do so, let $M = \frac{\alpha}{1-\alpha} 2 \sup_{\nu \in \mathcal{U}} |G(U)|$ and $\delta = (1 + K)\alpha)$. Since $K < (1-\alpha)/\alpha$, we have $\delta < 1$. Let T denote the first time such that $KM\delta^T < \varepsilon$. Note that for all $t < T$, we have $\max\{KM\delta^t, \varepsilon\} = KM\delta^t$.

We first show that for all $t < T$, we have $|U(c^t) - U(\tilde{c}^t)| \leq M\delta^t$. The proof works by induction. For $t = 0$, we have $c_0^t = \tilde{c}_0^t$, so

$$|U(c^0) - U(\tilde{c}^0)| = \sum_{s=1}^{\infty} \alpha^s |G(U({}_s c^0)) - G(U({}_s \tilde{c}^0))| \leq M$$

Suppose the claim holds for $t < T - 1$, we will show it holds for $t + 1$. We have

$$\begin{aligned} |U(c^{t+1}) - U(\tilde{c}^{t+1})| \leq\ & \alpha |G(U({}_1 c^{t+1})) - G(U({}_1 \tilde{c}^{t+1}))| \\ & + \alpha \sum_{s=1}^{\infty} \alpha^s |G(U({}_{s+1} c^{t+1})) - G(U({}_{s+1} \tilde{c}^{t+1}))|. \end{aligned} \tag{16}$$

By the induction hypothesis, the sum in (16) is bounded above by $M\delta^t$. And because $t < T - 1$, we have $KM\delta^t \geq \varepsilon$. Therefore,

$$|U(c) - U(c')| \leq \alpha KM\delta^t + \alpha M\delta^t \leq M\delta^{t+1},$$

which shows the claim.

Finally, for $t = T$, (16) still applies, but this time the first term is bounded by $\alpha\varepsilon$, because $KM\delta^T < \varepsilon$. This implies that

$$|U(c) - U(c')| \leq \alpha\varepsilon + \alpha M\delta^T \leq \alpha\varepsilon + \alpha\varepsilon/K = \delta\varepsilon/K.$$

Since $\delta < 1$ and $K > (1-\alpha)/2\alpha$, (A.3) follows.

Part (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 $\mathcal{J}$ be the operator on $B(M)$ defined by

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

By construction, $\mathcal{J}(U)$ 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 $K < (1-\alpha)/\alpha$, $\mathcal{J}$ must be a contraction, as is easily checked. So, $\mathcal{J}$ 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 U^* to (5) over $C(B) = \cup_M C(M)$.

Let $\mathcal{H}$ be the set of H-continuous functions. To verify that $U^* \in \mathcal{H}$, it suffices to show that (a) $\mathcal{J}$ maps $\mathcal{H}$ onto itself, and (b) $\mathcal{H}$ is closed under the sup norm. Indeed, this will guarantee that $\mathcal{J}$'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 $\frac{\alpha^T 2\bar{G}}{1-\alpha} < \varepsilon/2$, where $\bar{G} = \sup_{\nu\in\mathcal{U}} |G(\nu)|$. 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 $\tilde{c}$,

$$\begin{aligned} |\mathcal{J}(U)(c) - \mathcal{J}(U)(\tilde{c})| &\leq \left|\sum_{t=1}^{\infty} \alpha^t \left[G(U({}_tc)) - G(U({}_t\tilde{c}))\right]\right| \\ &\leq K\sum_{t=1}^{T-1} \alpha^t |U({}_tc) - U({}_t\tilde{c})| + \alpha^T \frac{2\bar{G}}{1-\alpha}. \end{aligned}$$

where K is the Lipschitz constant of G. The first term is less than $\frac{K\alpha}{(1-\alpha)} \max_{t\leq T-1} |U({}_tc) - U({}_t\tilde{c})|$. Now suppose that $c_t = \tilde{c}_t$ for all $t \leq N' = N + T$. This implies that $({}_tc)_{t'} = ({}_t\tilde{c})_{t'}$ for all $t \leq T$ and $t' \leq N$, because ${}_tc$ 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 $|U({}_tc) - U({}_t\tilde{c})| < \varepsilon/2$ for all $t \leq T$ and, hence, $|\mathcal{J}(U)(c) - \mathcal{J}(U)(\tilde{c})| < \varepsilon$. Setting $T(\varepsilon) = N'$ shows that $\mathcal{J}(U)$ satisfies H-continuity. To prove (b), consider a sequence $\{U^m\}$ 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, $|U^*(c^n) - U^*(c^m)| < \varepsilon$ for all $n, m \geq T$, since the consumption levels of c^n and 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 $\mathcal{J}$ on $C(B)$ and c^n belongs to $C(B)$, for each n

$$U^*(c^n) = u(c_0^n) + \sum_{t=1}^{\infty} \alpha^t G(U^*({}_t c^n))$$

The left-hand side converges to $U^*(c)$. Moreover, for each t, $U^*({}_t c^n)$ converges to $U^*({}_t c)$ (which is similarly well defined). Since G is continuous, $G(U^*({}_t c^n))$ converges to $G(U^*({}_t c))$ for each t. Since $\alpha < 1$ and G is bounded, by the dominated convergence theorem, the right-hand side converges to $u(c_0) + \sum_{t=1}^{\infty} \alpha^t G(U^*({}_t c))$, 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 $\mathcal{J}$ 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 ε was arbitrary, $U(c) = U^*(c)$ for all c, establishing uniqueness.

A.4 Proof of Proposition 2

Let $U(c) = V(c_0, U({}_1 c), U({}_2 c), \ldots)$ where V is strictly increasing in $U({}_t c)$ for all $t \geq 1$. By definition, $(x, c) \succ (y, c)$ means that $U(x, c) > U(y, c)$. Hence, for all $0 \leq s \leq t$,

$$U({}_s z_t, x, c) > U({}_s z_t, y, c),$$

where, for $s < t$, ${}_s z_t = (z_s, \ldots, z_t)$ and ${}_t z_t = z_t$. This follows by induction. For $s = t$,

$$U({}_t z_t, x, c) = V({}_t z_t, U(x, c), U(c), \ldots) > V({}_t z_t, U(y, c), U(c), \ldots) = U({}_t z_t, y, c).$$

Now suppose that the claim holds for $r + 1 \leq s \leq t$, with $0 \leq r < t$. Then

$$\begin{aligned} U({}_r z_t, x, c) &= V(z_r, U({}_{r+1} z_t, x, c), \ldots, U({}_t z_t, x, c), U(x, c), \ldots) \\ &> V(z_r, U({}_{r+1} z_t, y, c), \ldots, U({}_t z_t, y, c), U(y, c), \ldots) \;=\; U({}_r z_t, y, c). \end{aligned}$$

By definition, $({}_0 z_t, x, \hat{x}, c') \sim ({}_0 z_t, y, \hat{y}, c')$ means that

$$\begin{aligned} & V(z_0, U({}_1 z_t, x, \hat{x}, c'), \ldots, U({}_t z_t, x, \hat{x}, c'), U(x, \hat{x}, c'), U(\hat{x}, c'), \ldots) \\ = {} & V(z_0, U({}_1 z_t, y, \hat{y}, c'), \ldots, U({}_t z_t, y, \hat{y}, c'), U(y, \hat{y}, c'), U(\hat{y}, c'), \ldots). \end{aligned}$$

Since $U({}_0 z_t, x, c) > U({}_0 z_t, y, c)$ for all c,

$$V(z_0, U({}_1 z_t, y, \hat{x}, c'), \ldots, U({}_t z_t, y, \hat{x}, c'), U(y, \hat{x}, c'), U(\hat{x}, c'), \ldots)$$

$$< V(z_0, U({}_1z_t, y, \hat{y}, c'), \ldots, U(z_t, y, \hat{y}, c'), U(y, \hat{y}, c'), U(\hat{y}, c'), \ldots).$$

This implies that $U(\hat{y}, c') > U(\hat{x}, c')$. Otherwise, $U(y, \hat{y}, c') \leq U(y, \hat{x}, c')$ and, by induction, $U({}_sz_t, y, \hat{y}, c') \leq U({}_sz_t, y, \hat{x}, c')$ for all $0 \leq s \leq t$, which is a contradiction.

Finally, we must have $U(x, \hat{x}, c') > U(y, \hat{y}, c')$. Otherwise, again by induction, for all $0 \leq s \leq t$

$$U({}_sz_t, y, \hat{y}, c') > U({}_sz_t, x, \hat{x}, c'),$$

which contradicts $({}_0z_t, x, \hat{x}, c') \sim ({}_0z_t, y, \hat{y}, c')$.

Suppose that we replace condition $({}_0z_t, x, \hat{x}, c') \sim ({}_0z_t, y, \hat{y}, c')$ with $({}_0z_t, x, {}_{t+2}z_s, \hat{x}, c') \sim ({}_0z_t, y, {}_{t+2}z_s, \hat{y}, c')$ where $s \geq t+2$. By the same argument as before, $({}_0z_t, y, {}_{t+2}z_s, \hat{x}, c') \prec ({}_0z_t, y, {}_{t+2}z_s, \hat{y}, c')$ and so $(\hat{y}, c') \succ (\hat{x}, c')$. If not, by induction $({}_\tau z_s, \hat{x}, c') \succsim ({}_\tau z_s, \hat{y}, c')$ for all $0 \leq \tau \leq s$ (where $z_{t+1} = y$). Then, we must have $(x, {}_{t+2}z_s, \hat{x}, c') \succ (y, {}_{t+2}z_s, \hat{y}, c')$. If not, since $({}_\tau z_s, \hat{y}, c') \succ ({}_\tau z_s, \hat{x}, c')$ for all $t+2 \leq \tau \leq s$, we would have $({}_0z_t, y, {}_{t+2}z_s, \hat{y}, c') \succ ({}_0z_t, x, {}_{t+2}z_s, \hat{x}, c')$.

A.5 Proof of Theorem 4

Using Axiom 10 and Theorem 3, we also have

$$(c_0, c_1, {}_2c) \succ (c_0, c_1', {}_2c) \Leftrightarrow (\hat{c}_0, c_1, {}_2c') \succ (\hat{c}_0, c_1', {}_2c') \tag{17}$$

$$(c_0, c_1, {}_2c) \succ (c_0, c_1, {}_2c') \Leftrightarrow (\hat{c}_0, c_1', {}_2c) \succ (\hat{c}_0, c_1', {}_2c') \tag{18}$$

$$(c_0, c_1, {}_2c) \succ (c_0', c_1, {}_2c) \Leftrightarrow (c_0, c_1', {}_2c') \succ (c_0', c_1', {}_2c') \tag{19}$$

$$(c_0, c_1, {}_2c) \succ (c_0, c_1', {}_2c') \Leftrightarrow (\hat{c}_0, c_1, {}_2c) \succ (\hat{c}_0, c_1', {}_2c') \tag{20}$$

By Debreu's (1960) Theorem 3, conditions (17)-(20) and (i)-(ii) in Axiom 10 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 3, $\succ$ is also represented by

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

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

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

for some $\xi > 0$ 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 1, $\gamma < \frac{1-\alpha}{\alpha}$.

A.6 Proof of Proposition 5

By assumption, for all t,

$$U({}_tc) = u(c_t) + \sum_{\tau=t+1}^{\infty} \beta\delta^{\tau-t}u(c_\tau), \tag{21}$$

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

For the "if part" see the main text. For the "only if" part, using (21), 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, $\sum_{t=0}^{\infty} w(t)U({}_tc) = \sum_{t=0}^{\infty}\delta^t u(c_t)$. So 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,

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

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

References

Abel, A. B. (1985). Precautionary saving and accidental bequests. *American Economic Review 75*, 777–791.

Akerlof, G. A. (1991). Procrastination and obedience. *American Economic Review 81*, 1–19.

Allais, M. (1953). Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école américaine. *Econometrica 21*, 503–546.

Andreoni, J. (1989). Giving with impure altruism: applications to charity and ricardian equivalence. *Journal of Political Economy 97*, 1447–1458.

Andreoni, J. and C. Sprenger (2012). Estimating time preferences from convex budgets. *American Economic Review 102*, 3333–3356.

Asheim, G. B. (2010). Intergenerational equity. *Annual Review of Economics 2*, 197–222.

Asheim, G. B., T. Mitra, and B. Tungodden (2012). Sustainable recursive social welfare functions. *Economic Theory 49*, 267–292.

Barro, R. J. (1974). Are government bonds net wealth? *Journal of Political Economy 82*, 1095–1117.

Basu, K. and T. Mitra (2003). Aggregating infinite utility streams with intergenerational equity: the impossibility of being paretian. *Econometrica 71*, 1557–1563.

Becker, G. S. and N. Tomes (1979). An equilibrium theory of the distribution of income and intergenerational mobility. *Journal of Political Economy 87*, 1153–1189.

Bergstrom, T. (1989). Puzzles: Love and spaghetti, the opportunity cost of virtue. *Journal of Economic Perspectives 3*, 165–173.

Bergstrom, T. (1997). A survey of theories of the family. *Handbook of population and family economics*.

Bergstrom, T. C. (1999). Systems of benevolent utility functions. *Journal of Public Economic Theory 1*, 71–100.

Bernheim, B. D. (1989). Intergenerational altruism, dynastic equilibria and social welfare. *Review of Economic Studies 56*, 119–128.

Bernheim, B. D. and A. Rangel (2007). Toward choice-theoretic foundations for behavioral welfare economics. *American Economic Review 97*, 464–470.

Bernheim, B. D. and A. Rangel (2009). Beyond revealed preference: Choice-theoretic foundations for behavioral welfare economics. *Quarterly Journal of Economics 124*, 51–104.

Bernheim, B. D., A. Shleifer, and L. H. Summers (1985). The strategic bequest motive. *Journal of Political Economy 93*, 1045–1076.

Chakraborty, A. and Y. Halevy (2016). Allais meets Strotz: remarks on the relation between present bias and the certainty effect. *Vancouver School of Economics*.

Dasgupta, P. (2008). Discounting climate change. *Journal of Risk and Uncertainty 37*, 141–169.

Dasgupta, P. and E. Maskin (2005). Uncertainty and hyperbolic discounting. *American Economic Review 95*, 1290–1299.

de Tocqueville, A. (1835). *Democracy in America*. Saunders and Otley, London.

Debreu, G. (1954). Representation of a preference ordering by a numerical function. *Decision Processes 3*, 159.

Debreu, G. (1960). Topological methods in cardinal utility theory. *Mathematical Methods in the Social Sciences*, 16–26.

Diamond, P. A. (1965). The evaluation of infinite utility streams. *Econometrica 33*, 170–177.

Echenique, F., T. Imai, and K. Saito (2016). Testable implications of models of intertemporal choice: Exponential discounting and its generalizations. *Mimeo, Caltech*.

Farhi, E. and I. Werning (2007). Inequality and social discounting. *Journal of Political Economy 115*, 365–402.

Farmer, J. D. and J. Geanakoplos (2009). Hyperbolic discounting is rational: Valuing the far future with uncertain discount rates. *Mimeo, Yale University*.

Fels, S. and R. Zeckhauser (2008). Perfect and total altruism across the generations. *Journal of Risk and Uncertainty 37*, 187–197.

Frederick, S. (2003). Time preference and personal identity. *Time and Decision: Economic and Psychological Perspectives of Intertemporal Choice 27*, 89–113.

Fudenberg, D. and J. Tirole (1991). Game theory, 1991. *Cambridge, Massachusetts*.

Galperti, S. and B. Strulovici (2014). From anticipations to present bias: A theory of forward-looking preferences. *Mimeo, Northwestern University, Available at SSRN: http://ssrn.com/abstract=2517742*.

Gul, F. and W. Pesendorfer (2001). Temptation and self-control. *Econometrica 69*, 1403–1435.

Halevy, Y. (2008). Strotz meets Allais: Diminishing impatience and the certainty effect. *American Economic Review 98*, 1145–1162.

Halevy, Y. (2015). Time consistency: Stationarity and time invariance. *Econometrica 83*, 335–352.

Hayashi, T. (2003). Quasi-stationary cardinal utility and present bias. *Journal of Economic Theory 112*, 343–352.

Kahneman, D. and A. Tversky (1979). Prospect theory: An analysis of decision under risk. *Econometrica 47*, 263–291.

Kimball, M. S. (1987). Making sense of two-sided altruism. *Journal of Monetary Economics 20*, 301–326.

Koopmans, T. C. (1960). Stationary ordinal utility and impatience. *Econometrica 28*, 287–309.

Koopmans, T. C. (1972). Representation of preference orderings over time. *Decision and Organization*, 79–100.

Koopmans, T. C., P. A. Diamond, and R. E. Williamson (1964). Stationary utility and time perspective. *Econometrica 32*, 82–100.

Kőszegi, B. and A. Szeidl (2012). A model of focusing in economic choice. *Quarterly Journal of Economics 128*, 53–104.

Laibson, D. (1997). Golden eggs and hyperbolic discounting. *Quarterly Journal of Economics 112*, 443–478.

Lipset, S. M. (1960). *Political man: The social bases of politics.* Garden City, New York: Anchor Books.

Loury, G. C. (1981). Intergenerational transfers and the distribution of earnings. *Econometrica 49*, 843–867.

O'Donoghue, T. and M. Rabin (1999). Doing it now or later. *American Economic Review 89*, 103–124.

Olea, J. and T. Strzalecki (2014). Axiomatization and measurement of quasi-hyperbolic discounting. *Quarterly Journal of Economics 129*, 1449–1499.

Parfit, D. (1971). Personal identity. *Philosophical Review 80*, 3–27.

Parfit, D. (1976). Lewis, Perry, and what matters. In A. Rorty (Ed.), *The Identities of Persons*, pp. 91–107. University of California Press.

Parfit, D. (1982). Personal identity and rationality. *Synthese 53*, 227–241.

Pearce, D. G. (2008). Nonpaternalistic sympathy and the inefficiency of consistent intertemporal plans. In M. O. Jackson and A. McLennan (Eds.), *Foundations in Microeconomic Theory*, pp. 213–231. Springer Berlin Heidelberg.

Peleg, B. and M. E. Yaari (1973). On the existence of a consistent course of action when tastes are changing. *Review of Economic Studies 40*, 391–401.

Phelan, C. (2006). Opportunity and social mobility. *Review of Economic Studies 73*, 487–504.

Phelps, E. S. and R. A. Pollak (1968). On second-best national saving and game-equilibrium growth. *Review of Economic Studies 35*, 185–199.

Ramsey, F. P. (1928). A mathematical theory of saving. *Economic Journal 38*, 543–559.

Rawls, J. (1971). *A Theory of Justice*. Harvard University Press.

Ray, D. (1987). Nonpaternalistic intergenerational altruism. *Journal of Economic Theory 41*, 112–132.

Ray, D. (2014). Hedonistic altruism and welfare economics. *NYU mimeo*.

Ray, D. and R. Wang (2001). On some implications of backward discounting. *Unpublished manuscript*.

Rogers, A. R. (1994). Evolution of time preference by natural selection. *American Economic Review 84*, 460–481.

Rubinstein, A. (2003). "Economics and psychology"? The case of hyperbolic discounting. *International Economic Review 44*, 1207–1216.

Saez-Marti, M. and J. W. Weibull (2005). Discounting and altruism to future decision-makers. *Journal of Economic theory 122*, 254–266.

Saito, K. (2011). Strotz meets Allais: Diminishing impatience and the certainty effect: Comment. *American Economic Review 101*, 2271–2275.

Saito, K. (2015). Impure altruism and impure selfishness. *Journal of Economic Theory 158*, 336–370.

Samuelson, P. A. (1937). A note on measurement of utility. *Review of Economic Studies 4*, 155–161.

Schneider, M. T., C. P. Traeger, and R. Winkler (2012). Trading off generations: Equity, discounting, and climate change. *European Economic Review 56*, 1621–1644.

Segal, U. (1992). Additively separable representations on non-convex sets. *Journal of Economic Theory 56*, 89–99.

Segal, U. (1994). A sufficient condition for additively separable functions. *Journal of Mathematical Economics 23*, 295–303.

Stern, N. (2007). *The economics of climate change: the Stern review.* Cambridge University Press.

Strotz, R. H. (1955). Myopia and inconsistency in dynamic utility maximization. *Review of Economic Studies 23*, 165–180.

Weitzman, M. (1999). Just keep discounting, but... In P. Portney and J. Weyant (Eds.), *Discounting and Intergenerational Equity*, pp. 23–30. Resources for the Future Washington, DC.