University of Zurich[UZH]

URPP Equality of Opportunity

Is it Double Counting, and Does it Matter?



Maya Eden

Equality of Opportunity Research Series #102
November 2026

URPP Equality of Opportunity

URPP Equality of Opportunity Discussion Paper Series No. 102, November 2026

Is it Double Counting, and Does it Matter?

Maya Eden
maya.eden@econ.uzh.ch
University of Zurich

The University Research Priority Program "Equality of Opportunity" studies economic and social changes that lead to inequality in society, the consequences of such inequalities, and public policies that foster greater equality of opportunity. We combine the expertise of researchers based at the University of Zurich's Faculty of Arts and Social Sciences, the Faculty of Business, Economics and Informatics, and the Faculty of Law.

Any opinions expressed in this paper are those of the author(s) and not those of the URPP. Research published in this series may include views on policy, but URPP takes no institutional policy positions.

URPP Discussion Papers often represent preliminary work and are circulated to encourage discussion. Citation of such a paper should account for its provisional character.

URPP Equality of Opportunity, University of Zurich, Schoenberggasse 1, 8001 Zurich, Switzerland
info@equality.uzh.ch, www.urpp-equality.uzh.ch

Is it Double Counting, and Does it Matter?

Maya Eden*

University of Zurich and CEPR

October 15, 2026

Abstract

Benefiting future generations may also benefit their ancestors who care about them. Should the utilitarian cost-benefit analysis take this into account, or would that be "double counting"? I show that the answer depends on one's theory of well-being, but that, for evaluating tradeoffs between the near and distant future, the quantitative stakes of this question are modest, at least within the Barro and Becker framework. When the economy's trajectory is common knowledge and intergenerational altruism is time-invariant, the verdicts of the utilitarian cost-benefit analysis are robust to the inclusion of altruistic benefits. Under some realistic departures from this benchmark, accounting for altruism raises the relative weight on the more distant future, but not by very much.

JEL Classification: E61, D64

Keywords: Social discounting, intergenerational altruism

*Email: maya.eden@econ.uzh.ch. I am grateful to Geir Asheim, Susanne Burri, Moritz Drupp, Luis Mota Freitas, Thomas Houlden, Felix Koch, Kevin Kuruc, Frikk Nesje, Florian Scheuer, Itai Sher and Ivan Werning for helpful comments, as well as for comments made by seminar participants at the Normative Economics and Economic Policy Webinar, the Center for Population Level Bioethics at Rutgers University, Oxford University, Toulouse School of Economics, Bocconi, Zurich ETH (KOF), The Society of Social Choice and Welfare Webinar, St. Gallen, Bristol, LMU, The Equality of Opportunity URPP Workshop on Inheritance, and the UZH Ethics Colloquium. I acknowledge support from the UZH URPP on Equality of Opportunity.

1 Introduction

The normative assessment of intergenerational tradeoffs is complicated by the presence of intergenerational altruism. For example, an argument can be made that climate change not only harms future generations but also impacts people today because they *worry* about its effects on their children's and grandchildren's lives.

Conceptually, it is not obvious how these worries should affect the utilitarian cost-benefit analysis. Are parents' concerns for their children *additional* reasons to combat climate change? Or is it enough to take into account the direct benefits to the children? The goal of this paper is to clarify this debate and to assess its quantitative stakes for the social willingness to pay for public investment, such as climate change mitigation.

I study this question within the classical utilitarian framework, in which all people matter equally, regardless of when they were born. In this framework, all that matters is each person's well-being, or "utility," and not how it is generated. The well-being that a person derives from anticipating his children's future is treated no differently from the well-being generated by his own circumstances. It is therefore necessary to understand the ways in which a person's well-being can depend on the circumstances of his descendants.

Different theories of well-being take different stances on this issue. However, prominent theories imply at least some utility derived through intergenerational altruism. Conceptually, then, ignoring intergenerational altruism may be a mistake. Quantitatively, however, I find that the implications of this mistake regarding the normative assessment of tradeoffs between the near and distant future are probably small, at least within the Barro and Becker framework.

I begin by deriving a theoretical benchmark in which intergenerational altruism has no effect on utilitarian cost-benefit analysis. The benchmark concerns comparisons of transparent outcome-and-belief profiles, in which all people correctly anticipate the trajectory of the economy. Intergenerational altruism is assumed to be time-invariant: the degree to which a person's well-being depends on the consumption utility of each of his descendants is the same in every generation. Under these assumptions, accounting for altruistic benefits has no effect on the verdicts of the utilitarian criterion. The reason is that altruism "cancels out." A person's altruism towards his children is offset by his parents' altruism towards him; his altruism towards his grandchildren is offset by his grandparents' altruism towards him; and so on. Intergenerational altruism therefore creates no asymmetry between generations. It simply scales all costs and benefits by a common constant.

I then consider two departures from this benchmark. The first relaxes the assumption that altruism is time-invariant. It turns out that, while the *level* of intergenerational altruism does not matter, *changes* in altruism across time do. In the standard Barro and Becker framework, the degree to which parents care about each of their children is decreasing in the number of children that they have. As family sizes shrink along the demographic transition, this model implies that altruism is rising over time. Accounting for intergenerational altruism therefore increases the relative benefits associated with improving the lives of later generations. Given realistic parameterizations, however,

the quantitative significance of this channel is modest, as the altruism parameters are, empirically, not very sensitive to family size.

The second departure concerns the assumption of transparency. A decision maker today cannot make the trajectory of the economy common knowledge, because some people are already deceased and their beliefs are immutable. People born later have more ancestors who care about them who are not yet deceased, and who may therefore benefit from anticipating their well-being. A plan to increase the consumption of someone in the distant future may please their (yet-to-be-born) ancestors, but increasing the consumption of an elderly person today, who has no living ancestors, entails no such indirect benefits.

This, too, tends to increase the relative importance of future consumption. However, in the Barro and Becker framework, this channel has only modest effects on the benefits of increasing the consumption of *young* people today relative to the benefits of increasing the consumption of future people. The reason is that young people, like future people, have living parents and grandparents. In the Barro and Becker framework, in which altruism decays exponentially across generations, the altruism of parents and grandparents accounts for the bulk of the indirect benefits. Consequently, recognizing the immutability of past beliefs has little effect on how we ought to evaluate tradeoffs between current young adults (the near future), and people born many generations from now (the distant future).

This paper is related to a large literature on social discounting (see Millner and Heal [2023] for a recent review). There are two main approaches to social discounting: the market-based approach, which is advocated by Nordhaus [2007], and the classical utilitarian approach, which is advocated by Stern [2008].[1] According to the market-based approach, one ought to discount benefits to future people at the market rate. In contrast, the utilitarian approach adopts a social welfare function that cares equally about current and future people.

In the market-based approach, public concerns about the well-being of future generations play a central role, because (ideally) these are reflected in long-term market interest rates.[2] The normative significance of altruistic preferences within the utilitarian framework is less obvious; clarifying it is the focus of this paper.

The rest of the paper is organized as follows. Section 2 summarizes the debate on accounting for altruism within the utilitarian framework, and places this paper's contribution within that literature. Section 3 presents the setup. Section 4 relates the framework to different theories of well-being and clarifies the altruistic benefits that should be accounted for within each of these theories. Section 5 introduces the social ordering. Section 6 establishes the theoretical benchmark, and

[1] Stern [2008] evokes classical utilitarianism in his argument for a zero social rate of time preference – that is, a social welfare function that cares equally about consumption utility today and consumption utility in the distant future. However, he specifies a 0.1% annual social rate of time preference to account for the risk of extinction. This is consistent with classical utilitarianism in a richer model with uncertainty (in particular, regarding the survival of humanity).

[2] Nesje [2024] points out that the market interest rate may not be an adequate representation of the degree to which people today care about future generations, because it may neglect cross-dynastic altruism: that is, the degree to which one person cares about the well-being of another's descendants.

Sections 7 and 8 consider the two departures from it, together with their quantitative implications. Section 9 concludes. All proofs are in Appendix A.

2 The Debate

The debate regarding how to account for other-regarding preferences in the utilitarian framework revolves around two issues. The first is foundational: are there ways in which others' outcomes may affect a person's own well-being and, if so, to what extent can these effects be inferred from people's behavior? The economics literature mostly focuses on the latter part of the question. Sen [1977], Dworkin [1990], Milgrom [1993], Hausman [2011], and Sher [2020] make the point that people's decisions may be motivated by moral concerns which are conceptually separate from their own self interest. As Hausman [2011] explains, people are able to distinguish between what is better *for them* and what is better *overall*, and their decisions may sometimes be driven by the latter.

This is the source of the double counting problem: if a person voluntarily engages in prosocial behavior, and does so for prosocial, rather that self-interested reasons, the revealed-preference approach might wrongfully conclude that he engages in this behavior because it is better *for him*. Thus, we would be overcounting the benefits from this activity, by taking the stance that it is not only good for the prosocial reasons but also for self-interested ones.

The risk of such double-counting is, by now, firmly understood. However, on its own, it does not provide an answer to the first part of the question: can there be *legitimate* ways in which others' outcomes affect one's well-being? If the answer is yes, then ignoring other-regarding preferences runs the risk of *under*counting. Section 4 analyzes this question through the lens of prominent theories of well-being, and shows that a blanket exclusion of utility derived through intergenerational altruism is difficult to justify. This is broadly in accordance with intuition: plausibly, the loss of a child is also (and perhaps even primarily) bad for the parents, rather than only for the child.[3]

The second issue concerns the problematic verdicts of the utilitarian criterion in cases in which people's well-beings depend on the circumstances of others. The most salient example concerns anti-social preferences (see Harsanyi [1977][4]). To demonstrate, assume that person A hates person B and wants person B to suffer. A utilitarian may choose to have person B suffer only because it entails the dubious benefit of person A's satisfaction. This is problematic.

Ray [2018] presents a similar example, more closely related to intergenerational altruism, which shows that the problems are not restricted to antisocial attitudes. Sally has a grandmother that cares deeply about her, but Ann does not. In this case, a utilitarian that accounts for Sally's grandmother's feelings would be tempted to give preferential treatment to Sally over Ann. This would be unfair.

[3] Bernheim [1989], Farhi and Werning [2007], Farhi and Werning [2013] and Ray [2018] consider models in which the utilitarian planner accounts for people's altruistic feeling towards their children. See also Kaplow [1995], Fleurbaey [2012], Decerf and der Linden [2016] and Treibich [2019], who take similar approaches in other environments with other-regarding preferences (not necessarily intergenerational altruism in particular).

[4] Page 30. Similar issues are also discussed in Dworkin [1990] and Goodin [1986].

These observations are often taken as grounds for "laundering preferences" by disregarding their other-regarding components (see Goodin [1986] and, more recently, Bierbrauer [2021]). However, some find this solution deeply unsatisfactory, because it misdiagnoses the problem (see Hart [1979][5]). The reason not to discriminate against Ann is not that Sally's grandmother's welfare doesn't *really* depend on what happens to Sally. Rather, we should refrain from discrimination because of a separate, non-welfarist commitment to equal treatment. Similarly, in the case of the sadist, the problem has to do with counting benefits which are obtained as a direct result of others' suffering — not a problem with other-regarding preferences per se. For example, it seems equally problematic to say that a rapist's enjoyment should count against his victim's suffering, even if it is not his victim's *suffering* that he enjoys. Laundering preferences is a problematic solution because it shoehorns these non-welfarist considerations into a welfarist framework, instead of tackling them head-on.[6]

The purpose of this paper is not to settle this debate, but rather to assess its stakes. I show that, in some cases, accounting for intergenerational altruism has no effect on the bottom line. In other cases, accounting for altruistic benefits may change the normative evaluation of tradeoffs between different generations. However, at least within the Barro and Becker framework, accounting for altruism has little effect on how we ought to evaluate tradeoffs between current young people and future generations (although it may have larger effects for how we evaluate tradeoffs involving currently-living older people).

3 Setup

Time is discrete and doubly infinite, indexed by $t \in \mathbb{Z}$. Let N_t denote the number of people born in period t. Throughout this paper, I take the sequence $\{N_t\}_{t=-\infty}^{\infty}$ as exogenous.[7]

Individuals from the same cohort have the same *all-things-considered utility*, which is denoted by the upper-case U_t. I will sometimes refer to U_t simply as the individual's *utility*. The all-things-considered utility is a cardinally measurable, interpersonally comparable quantity of well-being; as we will see, different theories of well-being imply different interpretations of U_t. Importantly, the all-things-considered utility, U_t, is not just an arbitrary *representation* of individual preferences (in fact, as I discuss in the next section, it may not be a representation of individual preferences at all).

The all-things-considered utility is distinct from the individual's *consumption utility*, denoted by the lower-case $u(t)$, which is the utility that the individual derives from his own consumption. The

[5] Particularly Section IV.

[6] Sher [2024] makes a similar point in the context of non-welfarist considerations relating to the design of tax policies.

[7] It is not obvious how to extend the utilitarian framework to variable population settings: for example, should the social goal be to maximize the sum of utilities or the average utility? These two stances coincide in fixed-population settings, but imply different implications in variable population settings (for example, they would imply different judgements regarding whether it is desirable to add a person whose utility is positive but below average). This challenge is beyond the scope of this paper. However, in Appendix B I discuss how the paper's results apply to a particular variable-population criterion, which is total utilitarianism.

two may differ because an individual's all-things-considered utility can depend on the consumption utilities of others, as well as on his beliefs and higher-order beliefs about them.

To formalize this dependence, for $t' \in \mathbb{Z}$, let $u(t, t')$ denote the *expected value* that generation t' assigns to the consumption utility of generation t (that is, the mean of generation t''s belief about $u(t)$ given its uncertainty). Generalizing, for $\vec{m} = (m_1, ..., m_k) \in \mathbb{Z}^k$, let $u(\vec{m}) = u(m_1, ..., m_k)$ denote the expected value that generation m_k assigns to $u(m_1, ..., m_{k-1})$.

Similarly, let U_t denote generation t's all-things-considered utility; $U_{t,\tau}$ denote generation τ's expectation of U_t; etc.

An *outcome-and-belief profile* is an infinite-dimensional object of the form

$$\mathbf{u} = \{\{u(\vec{m})\}_{\vec{m}\in\mathbb{Z}^k}\}_{k=1}^{\infty},$$

indicating each generation's consumption utility as well as the means of its beliefs and higher-order beliefs about other generations' consumption utilities.

The Barro and Becker framework. To develop an intuition for how beliefs and higher-order beliefs may enter the all-things-considered utility, it is useful to consider the Barro and Becker framework (Barro and Becker [1988], Barro and Becker [1989]). This framework will also be used for illustrating the main results, as well as for the quantitative analysis.

Barro and Becker specify the all-things-considered utility of generation t as

$$U_t = u(t) + \beta \left(\frac{N_{t+n}}{N_t}\right)^{1-\epsilon} U_{t+n} = \sum_{\tau=0}^{\infty} \left(\beta^{\tau} \left(\frac{N_{t+n\tau}}{N_t}\right)^{-\epsilon}\right) \left(\frac{N_{t+n\tau}}{N_t}\right) u(t+n\tau) \tag{1}$$

(where the last equality comes from substituting $U_{t+n} = u(t+n) + \beta \left(\frac{N_{t+2n}}{N_{t+n}}\right)^{1-\epsilon} U_{t+2n}$, and then substituting $U_{t+2n} = u(t+2n) + \beta \left(\frac{N_{t+3n}}{N_{t+2n}}\right)^{1-\epsilon} U_{t+3n}$, and so on).

In this framework, n is the number of periods between two consecutive generations. Each generation cares about their own consumption utility, $u(t)$, and the all-things-considered utility of its direct descendants. The parameter $\beta \in (0, 1)$ captures the relative weight of the two components, in the constant population benchmark in which $N_t = N_{t+n}$. The parameter $\epsilon \in (0, 1)$ captures the degree to which the altruistic benefits from each child are decreasing in the number of kids: if $\epsilon = 0$, a unit increase in $u(t+n)$, applied to each of t's direct descendants, generates a benefit of β times the number of descendants (N_{t+n}/N_t). When $\epsilon > 0$, the benefit increases less-than-proportionately with the number of kids.

Barro and Becker do not distinguish between generation t's beliefs about U_{t+n}, and the actual value of U_{t+n}. However, as we will see, this distinction is important. If we adopt a hedonistic interpretation of well-being, a person's utility can depend only on his own experiences and beliefs, and not what actually happens to other people.

A hedonistic interpretation of the Barro and Becker model would therefore specify

$$U_t = u(t) + \beta \left(\frac{N_{t+n}}{N_t} \right)^{1-\epsilon} U_{t+n,t}$$

In this model, what matters is generation t's beliefs about the next generation's all-things-considered utility, $U_{t+n,t}$ (rather than simply U_{t+n}). It is further assumed that only the mean of the belief matters, and not other moments of the belief distribution.

As expectations are linear, the expected value that t assigns to U_{t+n} can be rewritten as

$$U_{t+n,t} = u(t+n,t) + \beta \left(\frac{N_{t+2n}}{N_{t+n}} \right)^{1-\epsilon} U_{t+2n,t+n,t}$$

That is, the weighted sum of the expected value of $(t+n)$'s consumption utility and his belief about $(t+n)$'s belief about U_{t+2n} (the term $U_{t+2n,t+n,t}$). Iterating forward, we obtain

$$U_t = u(t) + \sum_{\tau=1}^{\infty} \left(\beta^{\tau} \left(\frac{N_{t+n\tau}}{N_t} \right)^{-\epsilon} \right) \left(\frac{N_{t+n\tau}}{N_t} \right) u(t+n\tau, ..., t) \tag{2}$$

That is, t's all-things-considered utility depends on his own consumption utility, his beliefs about his children's consumption utility, his beliefs about his children's beliefs about his grandchildren's consumption utility, etc.

Generalization. I generalize this model by assuming that the all-things-considered utility, U_t, is some linear combination of the elements of $\mathbf{u}$:

$$U_t(\mathbf{u}) = \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k} \varphi_{t,\vec{m}} \left(\frac{N_{m_1}}{N_t} \right) u(\vec{m}), \text{ where } \varphi_{t,t} = 1 \text{ and } \varphi_{t,\vec{m}} \geq 0 \tag{3}$$

Note that the standard definition of sums of form $\sum_{\mathbb{Z}^k}$ (for an integer k) is that, (a) regardless of the order of summation, the limit exists, and that (b) the limit is the same regardless of the order of summation.

The parameter $\varphi_{t,t}$ denotes the dependence of generation t's all-things-considered utility on its own consumption utility. This functional form also allows the all-things-considered utility to depend on the consumption utilities of other generations (with coefficients $\varphi_{t,t'}$), on beliefs about these consumption utilities (with coefficients $\varphi_{t,(t',t)}$), on others' beliefs about consumption utilities (with coefficients $\varphi_{t,(t',t'')}$), and on various higher-order beliefs. I restrict attention to cases in which altruism parameters are weakly positive ($\varphi_{t,\vec{m}} \geq 0$).

The normalization $\varphi_{t,t} = 1$ is a substantive assumption about interpersonal comparability: it reflects the stance that a unit increase in the utility that a person derives from his own consumption ($u(t) \mapsto u(t) + 1$) generates a unit increase in his all-things-considered utility ($U_t \mapsto U_t + 1$) —

regardless of the other sources of utility that he may have.[8][9] The plausibility of this assumption will be revisited in the following section, when we discuss different interpretations of well-being.

While $\varphi_{t,t}$ is restricted to be the same for all t, the coefficients $\varphi_{t,t+n}$ are not. For example, in the Barro and Becker framework, we have that $\varphi_{t,t+n} = \beta \left(\frac{N_{t+n}}{N_t}\right)^{-\epsilon}$; when $\epsilon > 0$, these coefficients depend on the relative sizes of cohorts t and $t+n$.

This framework can accommodate both forward-looking altruism (towards one's descendants) and backward-looking altruism (towards one's ancestors).[10] It can accommodate "non-paternalistic" or ("pure") forms of altruism, in which a person's all-things-considered utility depends on his descendants' all-things-considered utilities (U_t) or his beliefs about these objects, rather than directly on their consumption utilities; the Barro and Becker model is an example.[11]

The framework can also accommodate certain forms of paternalistic (or impure) altruism. For example, the specification $\varphi_{t,t} = \varphi_{t,t+n} = 1$ (and $\varphi_{\vec{m}} = 0$ otherwise) captures a paternalistic model in which parents care only about their children's consumption utilities, and not about the utilities that they derive from their own children's consumption utilities.

However, there are also certain forms of altruism that this framework cannot accommodate. The framework requires that all generations use the same notion of consumption utility, $u(t)$, to evaluate the consumption of generation t. This rules out certain forms of paternalistic altruism, in which parents care about their children's consumption choices in a way that is inconsistent with the objective of maximizing their consumption utilities. Warm glow preferences, in which parents care about their own *contributions* to their children's consumption, also cannot be captured here, because there is no distinction between consumption utility derived from one's parents' sacrifices and consumption utility derived from other sources.[12]

4 Theories of Well-Being

How should we interpret the utilities, $\{U_t\}$? This is an open normative question, and one that is crucial for the interpretation of the parameters $\{\varphi_{t,\vec{m}}\}$. The philosophical literature distinguishes between three classes of theories of well-being (see Crisp [2021]): hedonistic theories, desire-fulfillment

[8] The additivity assumption further rules out any interactions between the different sources of utility: for example, the utility that I derive from my beliefs about my child's consumption is completely independent from the utility that I derive from his actual consumption, as well as from the utility that I derive from my grandchild's consumption. Certain aspects of this additivity condition are more plausible than others. However, at the very least, linearity can be motivated as a first-order approximation.

[9] It is useful to contrast this interpersonal comparability assumption with the one made in Farhi and Werning [2013]. There, it is assumed that two people with the same income have the same utility, even if one of them spends more of that income on their children. Here, I am assuming instead that the quantity of utility that one derives from their own consumption is independent from any utility that one derives from their children.

[10] The class of preferences with linear two-sided altruism characterized by Kimball [1987], and the examples in section III in Hori and Kanaya [1989] ("Linear Aggregators"), fit within this framework. See also section 3.1 ("Additively Separate Interdependent Utility Functions") in Bergstrom [1999].

[11] See Saez-Marti and Weibull [2005], as well as the characterization of the quasi-hyperbolic model in Galperti and Strulovici [2017], which also fits within our framework.

[12] An example of intergenerational altruism with warm-glow preferences can be found in Andreoni [1989].

theories and objective list theories. Below, I show how these theories differ in their treatment of intergenerational altruism, and what each implies for the parameters in (3).

Hedonistic theories. According to hedonistic theories, a person's utility is a function of his mental states throughout his lifetime. In this class of theories, intergenerational altruism enters a person's utility insofar as it affects his mental states. We may feel happy for our children, or sad for them, and these feelings count towards our own well-being.

Given the exclusive focus on mental states, what matters is our beliefs about our children, and not what actually happens to them. For example, our hedonic utility would be affected by our worries about our children's future, or by the comfort of anticipating a bright future for them; however, it cannot be affected by what will actually happen to them after we die.

Hedonism implies that $\varphi_{t,t'} = 0$ for all $t' \neq t$, because a person can directly experience only his own consumption utility, and not the consumption utility of others. Similarly, it implies that $\varphi_{t,(m_1,...,m_k)} = 0$ for all $m_k \neq t$, because a person can experience only his own beliefs, and not the beliefs of others. The specification in (2) is an example of a hedonistic theory.

Within this class of theories, the interpersonal comparability assumption that $\varphi_{t,t} = 1$ for all t means that any two people with the same consumption will derive the same sensations from that consumption, regardless of their family structure. Under this assumption, the enjoyment of ice-cream is not accentuated or diminished by the happiness that one has from anticipating the well-being of their children. Similarly, the suffering associated with hunger or deprivation is the same for all people, regardless of their beliefs about the prospects of their descendants.

Under this interpretation, the parameter $\varphi_{t,t}$ captures how the individual's *happiness* depends on his own consumption utility; the parameters $\{\varphi_{t,(t',t)}\}_{t \neq t'}$ capture how his happiness depends on his beliefs about others' consumption; the parameters $\{\varphi_{t,(t',t'',t)}\}_{t',t''=-\infty}^{\infty}$ capture the happiness that he derives from his beliefs about others' beliefs about others' consumption; and so on.

Given a hedonistic interpretation of utilities, these parameters should be calibrated to capture how a person's happiness depends on his beliefs about the well-being of others. This is generally not the same as his willingness to trade off his own consumption against theirs, a point emphasized by Sen (1977), Dworkin (1990), Milgrom (1993), Hausman (2011) and Sher (2020). A person might feel morally responsible for providing for his children, and this sense of obligation might lead him to sacrifice more of his own happiness than his altruistic feelings alone would warrant. Estimates of intergenerational altruism based on revealed preferences — like the ones that will be used in this paper — should therefore be interpreted as upper bounds on the degree of altruism that is relevant for computing hedonic utilities.

Desire-fulfillment theories. According to desire-fulfillment theories, a person's utility depends on the extent to which his desires are actually fulfilled, and not merely on his subjective experiences. For example, if an author desires to write an impactful book, his utility should be evaluated based on whether his book was actually impactful, and not merely on whether he believed it to be. What matters is the objective success of one's life projects.

Within this class of theories, there is a nontrivial question regarding how to account for other-regarding preferences. Parfit (1987) considers an example in which he meets a stranger on a bus, who tells him about an ongoing fatal illness.[13] Parfit feels sympathy towards the stranger, and genuinely wants him to recover. However, it would be odd to say that the stranger's recovery would be "good for Parfit," even though it would fulfill this desire.

Parfit uses this example to argue that a person's utility can depend only on the fulfillment of his desires with respect to his own life.[14] He calls this theory "Success Theory." However, even within the confounds of this theory, he recognizes important ways in which one's well-being may depend on the lives of their children (page 495):

> Suppose that I try to give my children a good start in life. I try to give them the right education, good habits, and psychological strength. Once again, I am now an exile, and I shall never be able to learn what happens to my children. Suppose that, unknown to me, my children's lives go badly. One finds that the education that I gave him makes him unempolyable, another has a mental breakdown, another becomes a petty thief. If my children's lives fail in these ways, and these failures are in part the result of mistakes I made as their parent, these failures in my children's lives would be judged on the Success Theory to be bad for me. One of my strongest desires was to be a successful parent. What is now happening to my children, though it is unknown to me, shows that this desire is not fulfilled. My life failed in one of the ways in which I most wanted it to succeed. Though I do not know this fact, it is bad for me, and makes it true that I have had a worse life."

Success Theory is only one example of a desire-fulfillment theory, and a very restrictive one. If one of his children were to die in an accident unbeknown to him, this would not count as something that would be bad for the exiled Parfit, even if he invested a great deal in his children and cared deeply about their lives. While different variants of desire-fulfillment theory would assess this situation differently, this passage illustrates nicely why it is difficult to justify a blanket exclusion of intergenerational altruism within this framework. The desire we have for our children's lives to go well is one of our strongest. Disregarding it entirely would be unprincipled, to the point of undermining the theory itself.

I now turn to the desire-fulfillment interpretation of U_t in equation (3). Just as what matters for the author is whether or not their book is actually impactful, what matters for the parent is whether their child actually has a good life. This means that a person's utility may depend

[13] See Appendix I of Parfit (1987), page 494.

[14] It is not clear how best to square this restriction with Parfit's arguments about personal identity. In Part III of *Reasons and Persons*, Parfit (1987) argues that what determines the appropriate degree of rational self-concern is the extent to which we are psychologically connected to earlier or later versions of ourselves, not the fact that we are numerically identical to earlier or later versions of ourselves. In light of this, a young person's desire to become a doctor is — by Parfit's own lights — of very limited significance for that same person once she is older, at least if she is only weakly psychologically connected to her younger self.

on the consumption that his descendants actually enjoy, and not only on his beliefs about their consumption. This is consistent, for example, with the Barro and Becker model in (1).

The functional form in (3) allows for such dependence by specifying $\varphi_{t,t'} > 0$: that is, a person benefits directly from the consumption utility of his descendants, regardless of his awareness. One implication of this approach is that increasing our own consumption today also contributes, indirectly, to fulfilling the aspirations of our deceased ancestors, who would have wanted us to prosper.

Under a desire-fulfillment conception of utility, the parameter $\varphi_{t,\vec{m}}$ captures the degree to which generation t cares about the higher-order belief $u(\vec{m})$. For example, if people desire their children's happiness, and their children's happiness depends on their beliefs about their own children's consumption utility, then $\varphi_{t,(t+2n,t+n)}$ captures the extent to which generation t cares about its children's beliefs (generation $t+n$) about their grandchildren's consumption utility (generation $t+2n$).

As with hedonism, calibrating the altruism parameters demands caution: estimates based on revealed preferences might be too high, because people might feel morally obligated to take their children's interests into account even when doing so conflicts with their own desires.

In desire-fulfillment theories, interpersonal comparability is generally a problem. The ordinal preferences of each person provide a ranking of outcomes for that person; they do not, in general, provide a straightforward way to compare the outcomes of two different individuals. The normalization $\varphi_{t,t} = 1$ reflects the stance that the "degree" to which all individuals desire their own consumption is the same. An ice cream for me contributes just as much to my desire-fulfillment as an ice cream to you contributes to your desire-fulfillment. According to this interpretation, people with more children simply have more desires.

Objective list theories. Objective list theories are based on the idea that there is some objective list of characteristics that amount to a good life. For example, Fletcher proposes an objective list consisting of "Achievement, friendship, happiness, pleasure, self-respect and virtue" (Fletcher [2015]). Other scholars include different characteristics on the list, but the commonality of these approaches is the joint commitment to objectivity and pluralism: there are several, objective features that jointly determine one's well-being.

Within this class of theories, the question of how to account for intergenerational altruism depends on what is included on the list. For example, Fletcher's list includes "happiness", which, as discussed above, may depend on the well-being of one's children. Parfit's list includes, explicitly, "having children and being a good parent." However, there are also lists that are consistent with the view that one's well-being depends only on their own circumstances, and not on their family's. For example, Finnis's list is "Life, knowledge, play, aesthetic experience, sociability (friendship), practical reasonableness, "religion"" [15]

[15] All of the examples cited in this paragraph are taken from Fletcher [2015].

Equation (3) can accommodate such a list by specifying $\varphi_{t,t} = 1$ and $\varphi_{t,\vec{m}} = 0$ for all $\vec{m} \neq t$. I refer to this as the *narrow objective list theory*, as it implies that $U_t = u(t)$: a person's well-being depends only on certain characteristics of his own life, which are here summarized as the person's consumption utility ($u(t)$). Other objective list theories may imply other parameterizations; for example, if "happiness" is an item on the list, a person's beliefs about the well-being of their descendants might matter too ($\varphi_{t,(t+n,t)} > 0$).

It is worth clarifying that the interpretation of $u(t)$ as *consumption* utility is somewhat tenuous; it captures the (plausible) assumption that the objective list of desirable characteristics is easier to obtain when one has more resources at their disposal. To use Fletcher's list as an example, money cannot buy achievement, friendship, happiness, pleasure, self-respect or virtue. However, each of these may be more difficult to obtain when one is burdened by poverty.

Interpersonal comparability is less of a problem within objective list theories, given their "objective" nature. In this context, the assumption that $\varphi_{t,t} = 1$ plays a somewhat different role. An objective list theorist faces two decisions: what to include on the list, and how to aggregate the different items on the list. The assumption that $\varphi_{t,t} = 1$ is a restriction on both. It says that the utility that an individual derives directly from his own consumption should be an item on the list, and, further, that that item should enter additively into the index of well-being.

5 The Social Ordering

Equipped with a notion of individual utility, we still face the question of how to aggregate the utilities of different generations. In what follows, I focus on the classical utilitarian criterion, according to which all people matter equally, regardless of when they were born.

Formally, I assume a partial social ordering of outcome-and-belief profiles, given by the following utilitarian criterion:

$$\sum_{t=-\infty}^{\infty} N_t(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}})) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}} \tag{4}$$

One outcome-and-belief profile is better than another if the population-weighted sum of utility differences is positive. Note that this defines only a partial ordering: it is silent about pairs of outcome-and-belief profiles for which the infinite sum on the left is zero or undefined.[16]

5.1 Easily-comparable outcome-and-belief profiles

The results below concern pairs of outcome-and-belief profiles whose utility differences are summable, in the following sense.

[16] Utilitarianism is sometimes specified based on the social objective of maximizing the sum of utilities across all people who will ever live. However, in infinite-horizon problems, this criterion runs into convergence issues. The specification here can also run into convergence issues, but does so less often: for example, when population is constant and u_t and $\tilde{u}_t$ are both positive and constant, the sums $\sum_{t=-\infty}^{\infty} u_t$ and $\sum_{t=-\infty}^{\infty} \tilde{u}_t$ both equal ∞, and thus $\mathbf{u}$ and $\tilde{\mathbf{u}}$ cannot be ranked based on the sum of utilities. However, if $u_t > \tilde{u}_t$ for all t, the sum $\sum_{t=-\infty}^{\infty}(u_t - \tilde{u}_t)$ is positive, and hence $\mathbf{u}$ and $\tilde{\mathbf{u}}$ can be ranked according to the above criterion.

Definition. *Two outcome-and-belief profiles* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ *are* easily comparable *if*

$$\sum_{t=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k} \varphi_{t,\vec{m}} N_{m_1} \left| u(\vec{m}) - \tilde{u}(\vec{m}) \right| < \infty$$

When two outcome-and-belief profiles are easily comparable, the sum in (4) is well-defined, and its terms can be rearranged and regrouped freely (by Fubini's theorem for infinite series); both properties are used in the proofs. Note that easy comparability is a property of a *pair* of outcome-and-belief profiles, rather than of the outcome-and-belief profiles themselves. In particular, it does not require utilities to be bounded — only that the utility *differences* be summable.

6 Theoretical Benchmark

This section presents a theoretical benchmark based on the stylized assumption that the coefficients of intergenerational altruism are *time-invariant*, in the following sense.

Definition. *Intergenerational altruism is* time-invariant *if for every* $t, t' \in \mathbb{Z}$, $k \geq 1$ *and* $\vec{m} \in \mathbb{Z}^k$,

$$\varphi_{t,t+\vec{m}} = \varphi_{t',t'+\vec{m}}$$

where $t + (m_1, ..., m_k) = (t + m_1, ..., t + m_k)$.

Time-invariance means that the way a person's utility depends on the consumption utilities of his descendants and ancestors (and on various beliefs about them) is the same in every generation. For example, the extent to which my child's consumption utility affects my utility is the same as the extent to which my grandchild's consumption utility will affect my child's utility. In the Barro and Becker frameworks in (1) and (2), this assumption is satisfied when $\epsilon = 0$ or when fertility rates are constant ($N_{t+n}/N_t = N_n/N_0$ for all t).

The theoretical benchmark concerns the ranking of outcome-and-belief profiles that are *transparent* according to the definition below.

Definition. *An outcome-and-belief profile* $\mathbf{u}$ *is* transparent *if, for every* t, k, *and* $(m_1, ..., m_k) \in \mathbb{Z}^k$,

$$u(m_1, ..., m_k) = u(m_1)$$

In other words, an outcome-and-belief profile is transparent if all beliefs are correct, and the trajectory of the economy is common knowledge (in the sense that higher-order beliefs are correct as well). Transparency is a reasonable but onerous requirement. It is reasonable as a normative benchmark, in which policy is decided once and for all at time $t = -\infty$, and made common knowledge. This is an appealing normative benchmark because any improvement on it would have to exploit deviations from common knowledge.

Transparency is also an onerous requirement because, conditional on past beliefs, only one future is consistent with full transparency: the one in which those beliefs turn out to be correct.

Consequently, a decision maker at any time t never faces a choice between two distinct transparent outcome-and-belief profiles. Nonetheless, as we will see, the ranking of transparent outcome-and-belief profiles is a useful theoretical benchmark for understanding choices made at $t > -\infty$ as well.

The results in this paper require a technical assumption that rules out cases in which the welfare gains from a common-knowledge increase in any $u(t)$ are unbounded. To state it, define

$$a_t := \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t} \varphi_{\tau,\vec{m}}$$

The term a_t aggregates the various ways in which people care about the consumption utility of an individual from cohort t. To interpret it, consider a one-unit increase in $u(t)$ that is common knowledge (across all cohorts $\tau \in \mathbb{Z}$). The utility of each individual from cohort τ increases by

$$\sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t} \varphi_{\tau,\vec{m}} \left(\frac{N_t}{N_\tau}\right)$$

Summing across the N_τ individuals of each cohort τ, and then across cohorts, the aggregate welfare gain is

$$\sum_{\tau=-\infty}^{\infty} N_\tau \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t} \varphi_{\tau,\vec{m}} \left(\frac{N_t}{N_\tau}\right) = N_t \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t} \varphi_{\tau,\vec{m}} = N_t a_t$$

In the appendix, I establish that when altruism is time invariant, a_t is the same for all t.

Assumption 1. *$a_t < \infty$ for all t.*

This assumption means that a common-knowledge increase in $u(t)$ increases the sum of utilities across all people by a finite amount. I maintain this assumption throughout.

The theoretical benchmark is established in the following theorem.

Theorem 1. *Assume that intergenerational altruism is time-invariant and that $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are easily comparable and transparent. Then,*

$$\sum_{t=-\infty}^{\infty} N_t \left(u(t) - \tilde{u}(t)\right) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}}$$

This theorem establishes that, when intergenerational altruism is time-invariant, the social ranking of transparent outcome-and-belief profiles does not depend on the coefficients of intergenerational altruism. In particular, it does not matter whether people care a lot or a little about their descendants. Nor does it matter whether their utilities depend on the actual consumption utilities of their descendants, their beliefs about these consumption utilities, or exclusively on their

own consumption utilities. *All the different conceptions of welfare imply exactly the same social ordering of transparent outcome-and-belief profiles.*

To develop intuition for this result, consider the Barro and Becker specification in (1), and assume for simplicity that $N_t = 1$ for all t. Note that

$$\mathbf{u} \succ \tilde{\mathbf{u}} \Leftrightarrow \sum_{t=-\infty}^{\infty} \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}})\right) > 0$$

$$\Leftrightarrow \sum_{t=-\infty}^{\infty} \left(\sum_{\tau=0}^{\infty} \beta^\tau u(t+\tau n) - \sum_{\tau=0}^{\infty} \beta^\tau \tilde{u}(t+\tau n) \right) > 0$$

$$\Leftrightarrow \sum_{t=-\infty}^{\infty} \left(\sum_{\tau=0}^{\infty} \beta^\tau (u(t+\tau n) - \tilde{u}(t+\tau n)) \right) > 0$$

$$\Leftrightarrow \sum_{t=-\infty}^{\infty} \left(u(t) - \tilde{u}(t)\right) \left(1 + \beta + \beta^2 + ...\right) > 0$$

$$\Leftrightarrow \left(\sum_{\tau=0}^{\infty} \beta^\tau \right) \sum_{t=-\infty}^{\infty} \left(u(t) - \tilde{u}(t)\right) > 0$$

$$\Leftrightarrow \sum_{t=-\infty}^{\infty} \left(u(t) - \tilde{u}(t)\right) > 0$$

where the fourth line follows from rearranging the order of summation: the consumption utility of each cohort s is reached from exactly one cohort t at each dynastic distance τ (namely, $t = s - \tau n$), so that each difference $u(s) - \tilde{u}(s)$ is counted with total weight $1 + \beta + \beta^2 + ...$ (the rearrangement is valid because the outcome-and-belief profiles are easily comparable).

In particular, the coefficient of intergenerational altruism, β, does not matter for the ranking of these outcome-and-belief profiles.

More generally, time-invariant altruism does not matter because it cancels out. Our love for our children increases the welfare gains from raising their consumption; but our parents' love for us increases the welfare gains from raising our own consumption by just as much. When altruism is time-invariant, the two effects exactly offset each other, and altruism does not affect the assessment of tradeoffs between us and our children.

The benchmark rests on two assumptions. The first is that intergenerational altruism is time-invariant. As we will see, this assumption can fail in the standard Barro and Becker framework, in which parents' willingness to sacrifice their own consumption for the sake of any one of their children can be decreasing in the number of children they have. In that case, the altruism parameters rise over time as family sizes shrink, and altruism does not necessarily "cancel out."

The second assumption is that the outcome-and-belief profiles under consideration are transparent. This makes sense from the perspective of a decision maker at time $t = -\infty$, who can choose the outcome-and-belief profile of consumption utilities once and for all and let beliefs adjust

accordingly. It is far less natural for decision making at any time $t > -\infty$. An outcome-and-belief profile is transparent only if *all* beliefs, including those of deceased people, align with the realized consumption utilities. At time t, therefore, there can be at most one transparent outcome-and-belief profile: the one that conforms to the beliefs of past people (assuming their beliefs were consistent with each other). The time-t decision maker never gets to choose between two transparent outcome-and-belief profiles, and the benchmark therefore cannot be applied directly to decisions taken at time t.

The following sections consider how each of these assumptions might be relaxed for the purpose of guiding policy.

7 Time-Varying Altruism

When altruism parameters are time-invariant, a_t is time-invariant. In general, however, the ratios a_t/a_0 may vary across time. This implies the following generalization.

Theorem 2. *Let* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ *be easily comparable and transparent. Then,*

$$\sum_{t=-\infty}^{\infty} \left(\frac{a_t}{a_0}\right) N_t \left(u(t) - \tilde{u}(t)\right) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}}$$

This theorem establishes that the *levels* of intergenerational altruism do not matter for the social ranking; all that matters is how the altruism parameters *change* across time.

7.1 Quantitative implications

This section illustrates the quantitative implications of time-varying altruism within the Barro and Becker framework. Assume that the all-things-considered utility that an individual from cohort t derives from a transparent outcome-and-belief profile $\mathbf{u}$ is

$$U_t(\mathbf{u}) = \sum_{\tau=0}^{\infty} \underbrace{\left(\beta^{\tau}\left(\frac{N_{t+n\tau}}{N_t}\right)^{-\epsilon}\right)}_{\varphi(t,t+n\tau)+\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k}\varphi(t,(t+n\tau,\vec{m}))} \left(\frac{N_{t+n\tau}}{N_t}\right) u(t+n\tau)$$

Note that both (1) and (2) are consistent with this specification, as U_t is specified here only for transparent outcome-and-belief profiles, in which there is no distinction between beliefs and realizations.

Based on the above expression,

$$a_t = \sum_{\tau=0}^{\infty} \beta^{\tau}\left(\frac{N_t}{N_{t-n\tau}}\right)^{-\epsilon}$$

The value of a_t therefore depends on four inputs: the generational discount rate, β; the degree

Table 1: Calibration Parameters

Variable	Interpretation	Value(s)
n	Number of periods between consecutive generations	28 years
ϵ	Dependence of altruism on family size	0, 0.3
β	The weight assigned to the utility of direct descendants	0.3

to which altruism decreases with family size, ϵ; the age of parenthood, n; and the path of fertility rates, $\{N_{t+n\tau}/N_{t+n(\tau-1)}\}_{\tau=0}^{\infty}$.

Table 1 summarizes the parameters used for this calibration. I set the age difference between consecutive generations to be 28 years, which is the average age difference between mothers and their children.[17]

I set $\beta = 0.3$, as in Cordoba and Ripoll [2019]. Given the 28-year age difference between parents and children, this implies that at the time of having children, parents care about their children's consumption roughly as much as they care about their own (due to pure time discounting). In particular, if people apply an annual discount factor of $\delta = 0.96$ to their own consumption, then the weight that a person assigns to their own consumption at age 28 is $\delta^{28} = 0.3$, which is the weight that they assign to the consumption of each of their children (assuming a family with two children, so that ϵ plays no role).

As discussed in Sen [1977], the estimate $\beta = 0.3$ represents an upper bound, as it is based on parents' willingness to pay to support their children. It is possible that this willingness to pay reflects, in part, a moral obligation to look after one's children, or even a legal obligation to do so, rather than a "selfish" motivation to increase one's own utility. Setting β at a lower level would further reduce the altruistic benefits, which would only mitigate the implications of time-varying altruism.

In the macroeconomics literature, ϵ is typically calibrated based on the observation that households with more children tend to spend less on the consumption of each child. For example, Adhami et al. [2024] calibrate $\epsilon = 0.2$ based on analysis of household expenditure patterns documented in Lino [2011]. Cordoba [2015] and Cordoba and Ripoll [2019] calibrate $\epsilon = 0.3$ based on the relationship between health expenditure per child and family size, as documented in Dickie and Messman [2004].

These values are likely upper bounds, for at least two reasons. First, they do not account for potential increasing returns at the household level (Lino [2011]). For example, children can share a bedroom, and their food can be prepared in the same oven. This might be part of the reason why

[17] Data sources: UN, World Population Prospects (2024) – processed by Our World In Data, https://ourworldindata.org/grapher/births-by-age-of-mother.

Table 2: Births Per Woman

	1950	1978	2006	2100
World	4.9	3.8	2.6	1.8
Africa	6.5	6.6	5	2
Latin America	5.8	4.4	2.3	1.6
China	5.8	2.7	1.6	1.3
India	5.7	4.9	2.9	1.7
North America	3.1	1.8	2.0	1.6
Europe	2.7	2.0	1.5	1.6

Source: Historical estimates and projections from the UN World Population Prospects, processed by Our World in Data. Available at: https://ourworldindata.org/population-simulation-tool. Note that birth per person (the ratio N_{t+n}/N_t) is obtained from these data by dividing by 2.

parents spend less per child in larger families.

Second, these estimates ignore various issues of selection. Different parents face different quality-quantity tradeoffs, which leads them to select into different family sizes.[18] Intuitively, the parameter ϵ should be identified by examining how an *exogenous* change in the number of children would affect a given household's expenditure patterns.

Angrist et al. 2010 do exactly that, using a variety of natural experiments. For instance, having twins can be considered an exogenous shock to family size, especially in later pregnancies. Another example is the sex composition of the first two children: the authors show that having two children of the same gender can be thought of as an exogenous shock that increases the likelihood of having a third child. Using these sources of random variation, they find that additional siblings have no effect on child outcomes, such as educational attainment or income in adulthood. This suggests that parents' investment in each of their children does not depend on the number of children, which is broadly consistent with $\epsilon = 0$. For the purpose of this analysis, I consider $\epsilon = 0.3$ (as in Cordoba and Ripoll 2019) as an upper bound; recall that, for $\epsilon = 0$, altruism is time-invariant and we are back at the theoretical benchmark.

The sequence of fertility rates is estimated based on the data summarized in Table 2 (where N_{t+n}/N_t is the number of births per woman at time $t+n$ divided by 2, to obtain a measure of births per parent). For the purpose of this calibration, I assume that fertility rates prior to 1950

[18] See Doepke et al. 2024 for an overview of the income-fertility relationship and its change over time.

Table 3: Cross-Sectional Implications

	(1) $\frac{a_{2006}}{a_{1950}}$ $(\epsilon = 0)$	(2) $\frac{a_{2006}}{a_{1950}}$ $(\epsilon = 0.3)$	(3) $\frac{b_{2006}}{b_{1950}}$ $(\epsilon = 0)$	(4) $\frac{b_{2026}}{b_{1950}}$ $(\epsilon = 0.3)$
World	1	1.05	1.39	1.35
Africa	1	1.02	1.39	1.28
Latin America	1	1.08	1.39	1.36
China	1	1.12	1.39	1.41
India	1	1.05	1.39	1.33
North America	1	1.05	1.39	1.39
Europe	1	1.06	1.39	1.43

The numbers correspond to the weights that the social welfare function assigns to the consumption utilities of young adults, born in 2006, relative to the consumption utilities of their grandparents, born in 1950. The rows differ in the fertility rates used for computing the welfare weights. In the third and fourth columns, b_t is computed under the assumption that people born in 1950 are still alive but have no living ancestors, and people born in 2006 have living parents (born 1978) and grandparents (born 1950).

Table 4: Intertemporal Implications

	(1) $\frac{a_\infty}{a_{2006}}$ $(\epsilon = 0)$	(2) $\frac{a_\infty}{a_{2006}}$ $(\epsilon = 0.3)$	(3) $\frac{b_\infty}{b_{2006}}$ $(\epsilon = 0)$	(4) $\frac{b_\infty}{b_{2006}}$ $(\epsilon = 0.3)$
World	1	1.06	1.03	1.08
Africa	1	1.11	1.03	1.12
Latin America	1	1.07	1.03	1.09
China	1	1.05	1.06	1.08
India	1	1.08	1.03	1.1
North America	1	1.03	1.03	1.06
Europe	1	1.01	1.03	1.03

The numbers correspond to the weights that the social welfare function assigns to the consumption utilities of people born in the distant future, relative to the consumption utilities of current young adults, born in 2006. In columns (3) and (4), I assume that young adults have living parents (born 1978) and living grandparents (born 1950). Here, $a_\infty := \lim_{t\to\infty} a_t$, and $b_\infty := \lim_{t\to\infty} b_t$.

were at their 1950 level, and that long-run fertility rates are equal to their UN projections for the year 2100. This implies that all of the ancestors of current (older) people had large families (in accordance with the 1950 numbers), whereas essentially all of the ancestors of future people will have small families (in accordance with the 2100 numbers).

I consider two questions. The first concerns the cross-sectional distribution of the weights, a_t, between the current young (born in 2006: currently 20 years old) and the current old (born in 1950: currently 76 years old). These results are presented in the first two columns of Table 3. The second concerns intertemporal tradeoffs, between the "near future" — things that will affect the lives of current young adults — and the "distant future" — things that will affect the lives of people born many generations from now. These results are presented in the first two columns of Table 4.

Cross-sectional implications. The first column of Table 3 contains the relative weights on the consumption utilities of the current old and the current young, under the assumption of time-invariant altruism ($\epsilon = 0$). These are all 1, in line with Theorem 1. The second column repeats this computation under the assumption that altruism is time-varying, with $\epsilon = 0.3$ parameterizing the dependence of the altruism parameters on family size. These results indicate that time-varying altruism along the demographic transition increases the relative weights on the consumption utilities of young people, but that the effects are modest in most places. The largest effect is in China, where

fertility rates dropped dramatically from 5.8 in 1950 to 1.6 in 2006. Even there, time variation in intergenerational altruism increases the weight on the consumption utility of the current young only by 12% relative to the current old.

Intertemporal implications. The first two columns of Table 4 contain estimates of the weights on the consumption utilities of people born in the distant future, relative to people born in 2006. Again, under the assumption of time-invariant altruism, there is no difference, which explains why the first column is populated with 1's. This is the benchmark of Theorem 1. The second column considers time-varying altruism, by setting $\epsilon = 0.3$. Given that fertility rates are projected to decline, this increases the weights on the consumption of future people relative to the consumption of the current young. Again, the effects are modest. The largest effects are in places in which the demographic transition is still underway: for example, in Africa, fertility rates in 2006 were around 5, and they are projected to decline to around 2 in the long-run. As a result, the relative weight on the consumption utilities of people born in the distant future is 11% higher than the weight on the consumption utilities of the people born in 2006. In Europe, where the demographic transition has already been largely completed in 2006, the time-variation in altruism increases the weight on the consumption utilities of future people by only 1%.

These results demonstrate that even the upper-bounds, $\beta = 0.3$ and $\epsilon = 0.3$, generate time-variation in intergenerational altruism that is quite modest. More realistic estimates of β and ϵ, that are closer to 0, would imply even more modest effects. The implication is that, at least within the Barro and Becker model, time-variation in intergenerational altruism is unlikely to make a big difference for the utilitarian ranking of transparent outcome-and-belief profiles.

8 Departures from Common Knowledge

As discussed above, the ranking of transparent outcome-and-belief profiles is of limited use for guiding decisions at any time $t > -\infty$, when the beliefs of deceased people are already immutable. Before modifying the transparency requirement to better fit these circumstances, it is worth establishing the following corollary.

Corollary 1. *Assume that* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ *are easily comparable, but not necessarily transparent. Assume further that for every* $k \geq 1$ *and* $\vec{m} \in \mathbb{Z}^k$*, it holds that* $\varphi_{t,(t',\vec{m})} = 0$ *for all* t *and* t'*. Then,*

$$\sum_{t=-\infty}^{\infty} \left(\frac{a_t}{a_0}\right) N_t \left(u(t) - \tilde{u}(t)\right) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}}$$

This corollary extends Theorem 2 to cases in which beliefs may be false, but do not matter for utilities (consistent with a desire-fulfillment conception of utilities). It states that, in these cases as well, the levels of intergenerational altruism do not matter for the social ranking; only changes in the altruism parameters across time do (as in Theorem 2). For example, the social ranking

given the "narrow" objective list theory, $U_t = u(t)$, is the same as the social ranking given the desire-fulfillment interpretation of the Barro and Becker framework (as in (1)), provided that $\beta > 0$ and $\epsilon = 0$ (so that altruism is time-invariant).

This corollary is nearly immediate from the theorem: even if $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are not transparent, the restriction to all-things-considered utility functions that depend only on realizations and not on beliefs implies that beliefs do not matter. Hence, we can replace $\mathbf{u}$ and $\tilde{\mathbf{u}}$ with their transparent equivalents and apply the theorem.

Though simple, this corollary has an important implication. Suppose we endorse a desire-fulfillment notion of utility, and believe that (a) what people desire, beyond their own consumption, is the fulfillment of their descendants' desires (rather than, say, their descendants' *happiness*, which also depends on the beliefs they hold), and (b) the degree of intergenerational altruism is approximately constant over time. Then, *at any point in time*, we can solve for the optimal policy under the simplifying assumption that preferences are purely self-regarding. Under these assumptions, intergenerational altruism plays no role and can be ignored for the purpose of guiding policy.

Things are less clean when utilities depend not only on realizations but also on beliefs. Consider a decision maker at time t (the current period). Many people are already deceased at this point, and it is impossible to change their consumption utilities or their beliefs about current and future consumption.

Assume that, at time t, the oldest people alive are members of cohort $(t - A)$ (where $A \geq 0$). Consider the following weaker transparency condition, which constrains only the beliefs of people who are not yet deceased at time t.

Definition. *An outcome-and-belief profile* $\mathbf{u}$ *is* time-t transparent *if, for all* $k \geq 1$ *and all* $\vec{m} \in \mathbb{Z}^k$ *such that* $m_j \geq t - A$ *for all* j*, it holds that*

$$u(\vec{m}) = u(m_1)$$

Time-t transparency implies that the beliefs of current and future people correctly adjust to decisions taken today. However, it allows for the possibility that the beliefs of deceased people were mistaken. If past people believed that we would do one thing and we end up doing another, their beliefs will be falsified.

I further restrict attention to decision problems among alternatives that share the same past, as formally defined below.

Definition. *Two time-*t *transparent outcome-and-belief profiles* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ have the same past *if for all* $k \geq 1$ *and* $\vec{m} \in \mathbb{Z}^k$ *such that* $m_j < t - A$ *for some* $j \in \{1, ..., k\}$*, it holds that*

$$u(\vec{m}) = \tilde{u}(\vec{m})$$

This definition means that past consumption utilities are the same; that past beliefs are the same; and that future beliefs about past beliefs are the same (and, presumably, correct, although this is not required by this definition).

Consider a one-unit increase in $u(\tau)$ for some $\tau \geq t - A$, and assume that this increase is common knowledge among those individuals who are not yet deceased at time t. The extent to which this increase affects the utility of an individual from cohort τ' is

$$\hat{b}_{\tau',\tau} := \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k, \vec{m} \geq t-A, m_1 = \tau} \varphi_{\tau',\vec{m}} \left(\frac{N_\tau}{N_{\tau'}} \right)$$

Note that, even when we restrict attention to the comparison of time-t transparent outcome-and-belief profiles with the same past, deceased individuals may still be affected. If our theory of well-being assigns any role to desire fulfillment, this cannot be ruled out: for example, we might consider it to be better for a deceased individual to have a prospering dynasty, while not rejecting that his beliefs *also* play a (separate) role. In this case, we might say something like: it is better for our deceased ancestors that we prosper, even if they cannot be made aware of this. In addition, it may have been even better for them to have known during their lifetimes that we would be doing well.

The sum of welfare gains from such an increase across all cohorts is:

$$\sum_{\tau'=-\infty}^{\infty} N_{\tau'} \hat{b}_{\tau',\tau} = \sum_{\tau=-\infty}^{\infty} N_{\tau'} \left(\sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k, \vec{m} \geq t-A, m_1 = \tau} \varphi_{\tau',\vec{m}} \left(\frac{N_\tau}{N_{\tau'}} \right) \right) = N_\tau b_\tau, \text{ where}$$

$$b_\tau := \sum_{\tau'=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k, \vec{m} \geq t-A, m_1 = \tau} \varphi_{\tau',\vec{m}}$$

When altruism is time-invariant and $\varphi_{\tau',\vec{m}} \geq 0$ for all τ' and $\vec{m}$, the sequence $\{b_\tau\}_{\tau=t-A}^{\infty}$ is weakly increasing: later generations have longer lines of ancestors who are not yet deceased, and whose beliefs can still be affected. Note that backward-looking altruism creates no analogous asymmetry: every person who is not yet deceased has an infinite line of future descendants who are not yet deceased.

The following result provides an equivalent criterion for ranking two time-t transparent outcome-and-belief profiles that share the same past.

Theorem 3. *If* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ *are two easily comparable, time-t transparent outcome-and-belief profiles that share the same past, then*

$$\mathbf{u} \succ \tilde{\mathbf{u}} \Leftrightarrow \sum_{\tau=t-A}^{\infty} \left(\frac{b_\tau}{b_t} \right) N_\tau (u(\tau) - \tilde{u}(\tau)) > 0$$

This theorem suggests that intergenerational altruism may push policy in a more forward-looking direction, at least in cases in which the sequence $\{b_\tau\}$ is increasing (as is always the case with time-invariant altruism).

8.1 Quantitative implications

Fix some $t > -\infty$, and consider the following specification, in line with the hedonistic interpretation of the Barro and Becker model in (2):

$$\hat{b}_{\tau',\tau} = \begin{cases} \beta^{(\tau-\tau')/n} \left(\frac{N_\tau}{N_{\tau'}}\right)^{1-\epsilon} & \text{if } \tau \geq t - A \text{ and } \tau - \tau' = Tn \text{ for some } T \in \{0, 1, 2, ...\} \\ 0 & \text{otherwise.} \end{cases}$$

That is, an increase in $u(\tau)$ that is common knowledge among all those who are not yet deceased at time t benefits the individual himself ($T = 0$) and each of his direct ancestors T generations back, provided that they are alive at time t ($\tau' \geq t - A$); the ancestor's benefit is discounted by the generational discount factor, $\beta^T = \beta^{(\tau-\tau')/n}$, and adjusted for family size.

In this framework, once a person is deceased, there is no way to affect his utility. Hence, only the altruism of current and future people counts towards social welfare.

Cross-sectional implications. Columns (3) and (4) of Table 3 contain the ratios of the weights that a current decision maker (in 2026) would assign to the consumption utilities of the current young and the current old, given the hedonistic interpretation of the Barro and Becker model above and the parameters in Tables 1 and 2. The current young are assumed to have two generations of living ancestors (their parents and grandparents), and the current old — their grandparents — are assumed to have no living ancestors. This creates a large asymmetry between them: the weight on the consumption utility of the young is 28%-43% higher than the weight on the consumption utility of the old, depending on the choice of ϵ and on demographics.

Intertemporal implications. Columns (3) and (4) of Table 4 contain estimates of the weights on the consumption utilities of people born in the distant future, relative to people born in 2006, from the perspective of today (2026). Again, the hedonistic framework implies an asymmetry: the current young have only two generations of living ancestors who care about them, while future people have a long line of ancestors who care about them who are not yet deceased (or born). However, the effects of this asymmetry are quite modest.

To illustrate, consider the case of time-invariant altruism ($\epsilon = 0$) and note that, for $\beta = 0.3$,

$$1 + \beta + \beta^2 = 1.39 \text{ and } \sum_{t=0}^{\infty} \beta^t = 1.43 = 1.03(1 + \beta + \beta^2)$$

The first term is the value of b_{2006}, which includes the direct effect, 1; the indirect effect on the parents, β; and the indirect effect on the grandparents, β^2. The second term is the value of $\lim_{t\to\infty} b_t$, which includes the sum of all indirect effects on all ancestors. The similarity of these magnitudes suggests that the asymmetry between current young adults and future people is not

very large.[19]

Taken together, these results suggest that the immutability of past beliefs is unlikely to create a large asymmetry between the benefits of increasing consumption in the near future and the benefits of increasing consumption in the distant future. However, it creates a more substantial asymmetry between the benefits of increasing to consumption of the current young relative to the benefits of increasing the consumption of the current old.

9 Conclusion

Should utilitarian cost-benefit analysis count the benefits that improvements in future living standards confer on altruistic ancestors, or would doing so amount to "double counting"? This paper argues that the answer depends on one's theory of well-being. Under hedonism, a person's utility responds to his beliefs about his descendants' prospects; under desire-fulfillment theories, it responds to their actual consumption; only a narrow objective list theory, in which well-being depends solely on one's own circumstances, dismisses altruistic benefits outright. The blanket exclusion of intergenerational altruism from welfare analysis is therefore difficult to justify on conceptual grounds.

It can, however, be somewhat justified on quantitative grounds. In a benchmark in which the economy's trajectory is common knowledge and altruism is time-invariant, intergenerational altruism is irrelevant for the utilitarian ranking: a person's altruism towards his children is exactly offset by his parents' altruism towards him, so altruism merely scales all costs and benefits by a common constant, and the planner can proceed as if preferences were purely self-regarding. Realistic departures from this benchmark — rising altruism along the demographic transition, and the immutability of deceased people's beliefs — may alter some of these conclusions. However, a calibration of the Barro and Becker model suggests that accounting for altruism primarily affects the evaluation of cross-sectional tradeoffs between the current old and the current young. The evaluation of intertemporal tradeoffs between the near and distant future does not depend very much on how we account for intergenerational altruism.

[19] These calculations assume that the only beliefs that matter for people's utilities are their beliefs at the time of their death; in particular, they ignore the fact that the past beliefs of living people are also immutable, and that these beliefs may matter for their utilities. For example, we can only affect the beliefs that the grandparents of the current young adults hold in old age – not the beliefs that they held when they were younger. This assumption can be motivated by the following model. Assume that a person from generation τ' lives for A periods, indexed $a = 1, ..., A$. Let $u_a(\tau, \tau')$ denote the mean belief of τ' with respect to $u(\tau)$, at age a. Now assume that the altruism from anticipating the consumption utility of generation τ at age 1 is given by $\beta u_1(\tau, \tau')$ while, for any age $a > 1$, the hedonistic benefits are the *difference* $\beta u_a(\tau, \tau') - \beta u_{a-1}(\tau, \tau')$: that is, hedonistic utility responds to changes in beliefs. For example, I may be ecstatic to learn that my children will be ok if I previously expected them not to be, but I will not be similarly ecstatic if I already expected them to be ok. In this framework, the sum of hedonistic utilities from altruism throughout τ''s lifetime is given by $\beta u_1 + (\beta u_2 - \beta u_1) + ... + (\beta u_A - \beta u_{A-1}) = \beta u_A$; that is, the sum of utility derived from altruism is determined by end-of-life beliefs.

References

Mohamad Adhami, Mark Bils, Charles I. Jones, and Peter J. Klenow. Population and welfare: The greatest good for the greatest number. 2024.

James Andreoni. Giving with impure altruism: Applications to charity and ricardian equivalence. *Journal of political Economy*, 97(6):1447–1458, 1989.

Joshua Angrist, Victor Lavy, and Analia Schlosser. Multiple experiments for the causal link between the quantity and quality of children. *Journal of Labor Economics*, 28(4):773–824, 2010.

Robert J. Barro and Gary S. Becker. A reformulation of the economic theory of fertility. *The Quarterly Journal of Economics*, 103(1):1–25, 1988.

Robert J. Barro and Gary S. Becker. Fertility choice in a model of economic growth. *Econometrica*, 57(2):481–501, 1989.

Theodore C. Bergstrom. Systems of benevolent utility functions. *Journal of Public Economic Theory*, 1:71–100, 1999.

B. Douglas Bernheim. Intergenerational altruism, dynastic equilibria and social welfare. *The review of economic studies*, 56(1):119–128, 1989.

Felix Bierbrauer. Harry potter and the welfare of the willfully blinded. In Ralph Hertwig and Christoph Engel, editors, *Deliberate Ignorance: Choosing Not to Know*, pages 187–198. Cambridge, MA: MIT press, 2021.

Juan Carlos Cordoba. Children, dynastic altruism and the wealth of nations. *Review of Economic Dynamics*, 18(4), 2015.

Juan Carlos Cordoba and Marla Ripoll. The elasticity of intergenerational substitution, parental altruism, and fertility choice. *The Review of Economic Studies*, 86(5), 2019.

Roger Crisp. Well-being. In Edward N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2021. URL `https://plato.stanford.edu/archives/win2021/entries/well-being/`.

Benoit Decerf and Martin Van der Linden. Fair social orderings with other-regarding preferences. *Social Choice and Welfare*, 46:655–694, 2016.

Mark Dickie and Victoria L. Messman. Parental altruism and the value of avoiding acute illness: are kids worth more than parents? *Journal of Environmental Economics and Management*, 48 (3):1146–1174, 2004.

Matthias Doepke, Anne Hannusch, Fabian Kindermann, and Michéle Tertilt. The economics of fertility: A new era. *Handbook of the Economics of the Family*, 1:151–254, 2024.

Ronald Dworkin. The 'double counting' objection. In Jonathan Glover, editor, *Utilitarianism and Its Critics*, pages 103–106. MacMillan, 1990.

Emmanuel Farhi and Ivan Werning. Inequality and social discounting. *Journal of political economy*, 115(3):365–402, 2007.

Emmanuel Farhi and Ivan Werning. Estate taxation with altruism heterogeneity. *American Economic Review*, 103(3):489–495, 2013.

Guy Fletcher. 12 objective list theories. *The Routledge handbook of philosophy of well-being*, 2015.

Marc Fleurbaey. The importance of what people care about. *Politics, Philosophy and Economics*, 11:415–447, 2012.

Simone Galperti and Bruno Strulovici. A theory of intergenerational altruism. *Econometrica*, 85 (4):1175–1218, 2017.

Robert E Goodin. Laundering preferences. *Foundations of social choice theory*, 75:81, 1986.

John C Harsanyi. Rule utilitarianism and decision theory. *Erkenntnis*, 11(1):25–53, 1977.

Herbert LA Hart. Between utility and rights. *Columbia Law Review*, 79(5):828–846, 1979.

Daniel M Hausman. *Preference, value, choice, and welfare*. Cambridge University Press, 2011.

Hajime Hori and Sadao Kanaya. Utility functionals with nonpaternalistic intergenerational altruism. *Journal of Economic Theory*, 49:241–265, 1989.

Louis Kaplow. A note on subsidizing gifts. *Journal of Public Economics*, 58:469–477, 1995.

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

Mark Lino. Expenditures on children by families. Technical Report, U.S. Department of Agriculture, Center for Nutrition Policy and Promotion, Miscellaneous Publication No. 1528-2011, 2011.

Paul Milgrom. Is sympathy an economic value? philosophy, economics, and the contingent valuation method. In J. A. Hausman, editor, *Contingent Valuation: A Critical Assessment*, pages 417–441. Elsevier Science Publishers, 1993.

Antony Millner and Geoffrey Heal. Choosing the future: Markets, ethics, and rapprochement in social discounting. *Journal of Economic Literature*, 61(3):1037–87, September 2023.

Frikk Nesje. Cross-dynastic intergenerational altruism. *Journal of Environmental Economics and Management*, 128:103059, 2024.

William D. Nordhaus. A review of the stern review on the economics of climate change. *Journal of Economic Literature*, 45(3):686–702, 2007.

Derek Parfit. *Reasons and persons*. Oxford University Press, 1987.

Debraj Ray. Hedonistic altruism and welfare economics. 2018.

Maria Saez-Marti and Jörgen W. Weibull. Discounting and altruism to future decision-makers. *Journal of Economic Theory*, 122(2):254–266, 2005. ISSN 0022-0531. doi: https://doi.org/10.1016/j.jet.2004.06.003. URL https://www.sciencedirect.com/science/article/pii/S0022053104001358

Amartya K Sen. Rational fools: A critique of the behavioral foundations of economic theory. *Philosophy & public affairs*, pages 317–344, 1977.

Itai Sher. How perspective-based aggregation undermines the pareto principle. *Politics, Philosophy & Economics*, 19(2):182–205, 2020.

Itai Sher. Generalized social marginal welfare weights imply inconsistent comparisons of tax policies. *American Economic Review*, 114(11):3551–3577, 2024.

Nicholas Stern. The economics of climate change. *American Economic Review*, 98(2):1–37, 2008.

Rafael Treibich. Welfare egalitarianism with other-regarding preferences. *Social Choice and Welfare*, 52:1–28, 2019.

Stéphane Zuber and Geir B Asheim. Justifying social discounting: the rank-discounted utilitarian approach. *Journal of Economic Theory*, 147(4):1572–1601, 2012.

A Proofs

Throughout the appendix, $\mathbf{u}$ and $\tilde{\mathbf{u}}$ denote a pair of easily comparable outcome-and-belief profiles (as defined in Section 5), and I write $\Delta(\vec{m}) := u(\vec{m}) - \tilde{u}(\vec{m})$. By (3) and the linearity of U_t,

$$\sum_{t=-\infty}^{\infty} N_t \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}})\right) = \sum_{t=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k} \varphi_{t,\vec{m}} N_{m_1} \Delta(\vec{m}) \tag{5}$$

Easy comparability guarantees that this sum converges absolutely, so that its terms can be rearranged and grouped arbitrarily without affecting its value (by Fubini's theorem for infinite series). All of the results follow from rearranging (5), grouping terms according to the first coordinate of $\vec{m}$.

The following lemma is the key step. Because Theorem 1 is a special case of Theorem 2, I prove Theorem 2 first.

Lemma 1. *Assume that $0 < a_t < \infty$ for every t, and that $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are easily comparable and transparent. Then,*

$$\sum_{t=-\infty}^{\infty} N_t \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}}) \right) = \sum_{t=-\infty}^{\infty} a_t N_t \left(u(t) - \tilde{u}(t) \right)$$

Proof. Because $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are transparent, $u(\vec{m}) = u(m_1)$ and $\tilde{u}(\vec{m}) = \tilde{u}(m_1)$ for every $\vec{m}$, and hence $\Delta(\vec{m}) = \Delta(m_1)$. Substituting into (5) and grouping terms according to $s = m_1$ (which is permissible because $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are easily comparable),

$$\sum_{t=-\infty}^{\infty} N_t \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}}) \right) = \sum_{s=-\infty}^{\infty} N_s \Delta(s) \left(\sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k;\, m_1 = s} \varphi_{\tau,\vec{m}} \right) = \sum_{s=-\infty}^{\infty} a_s N_s \Delta(s)$$

□

Proof of Theorem 2. By Lemma 1 and $0 < a_0 < \infty$,

$$\sum_{t=-\infty}^{\infty} N_t \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}}) \right) = a_0 \sum_{t=-\infty}^{\infty} \left(\frac{a_t}{a_0} \right) N_t \left(u(t) - \tilde{u}(t) \right)$$

Because $a_0 > 0$, the left-hand side is strictly positive if and only if $\sum_{t=-\infty}^{\infty} (a_t/a_0) N_t \left(u(t) - \tilde{u}(t) \right) > 0$. The result then follows from the definition of the social ordering, (4). □

Proof of Theorem 1. It suffices to show that, when intergenerational altruism is time-invariant, $a_s = a$ for every s; the result then follows from Theorem 2, as $a_s/a_0 = 1$ for all s and $0 < a < \infty$ by assumption.

Time-invariance states that $\varphi_{t,t+\vec{m}} = \varphi_{t',t'+\vec{m}}$ for all t, t' and $\vec{m}$; equivalently, $\varphi_{\tau,\vec{m}} = \varphi_{0,\vec{m}-\tau}$ for all τ and $\vec{m}$, where $\vec{m} - \tau = (m_1 - \tau, ..., m_k - \tau)$. Hence, for every s,

$$a_s = \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k;\, m_1 = s} \varphi_{0,\vec{m}-\tau} = \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{v} \in \mathbb{Z}^k;\, v_1 = s-\tau} \varphi_{0,\vec{v}} = \sum_{j=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{v} \in \mathbb{Z}^k;\, v_1 = j} \varphi_{0,\vec{v}}$$

where the second equality substitutes $\vec{v} = \vec{m} - \tau$ (a bijection of $\mathbb{Z}^k$ for each fixed τ and k), and the third substitutes $j = s - \tau$ (a bijection of $\mathbb{Z}$ for each fixed s). The right-hand side does not depend on s, and equals a (take $s = 0$). □

Proof of Corollary 1. By assumption, $\varphi_{t,\vec{m}} = 0$ for every $\vec{m}$ of length $k \geq 2$. Hence, (3) reduces to

$$U_t(\mathbf{u}) = \sum_{t'=-\infty}^{\infty} \varphi_{t,t'} \left(\frac{N_{t'}}{N_t} \right) u(t')$$

so that $U_t(\mathbf{u})$ depends only on the consumption utilities $\{u(t')\}_{t'=-\infty}^{\infty}$, and not on beliefs or higher-order beliefs. Let $\hat{\mathbf{u}}$ denote the transparent counterpart of $\mathbf{u}$, defined by $\hat{u}(\vec{m}) = u(m_1)$ for every $\vec{m}$, and define $\hat{\tilde{\mathbf{u}}}$ analogously. Then $U_t(\hat{\mathbf{u}}) = U_t(\mathbf{u})$ and $U_t(\hat{\tilde{\mathbf{u}}}) = U_t(\tilde{\mathbf{u}})$ for every t, so that $\mathbf{u} \succ \tilde{\mathbf{u}}$

if and only if $\hat{\mathbf{u}} \succ \hat{\tilde{\mathbf{u}}}$. Moreover, the pair $(\hat{\mathbf{u}}, \hat{\tilde{\mathbf{u}}})$ is easily comparable (given the vanishing of the coefficients on all $\vec{m}$ of length $k \geq 2$, the relevant sums coincide with those of the original pair) and $\hat{u}(t) = u(t)$, $\hat{\tilde{u}}(t) = \tilde{u}(t)$ for every t. The result follows from applying Theorem 2 to $(\hat{\mathbf{u}}, \hat{\tilde{\mathbf{u}}})$. □

Proof of Theorem 3. Note that $b_\tau \leq a_\tau$ for all τ, and hence Assumption 1 guarantees that $b_\tau < \infty$.

Without loss of generality, I prove the theorem for $t = 0$ (that is, time-0 transparent profiles). Consider a vector $\vec{m} \in \mathbb{Z}^k$. If $m_j < -A$ for some $j \in \{1, ..., k\}$, then $\Delta(\vec{m}) = 0$, because $\mathbf{u}$ and $\tilde{\mathbf{u}}$ share the same past. Otherwise, $m_j \geq -A$ for all j, and time-0 transparency implies that $u(\vec{m}) = u(m_1)$ and $\tilde{u}(\vec{m}) = \tilde{u}(m_1)$, so that $\Delta(\vec{m}) = \Delta(m_1)$ with $m_1 \geq -A$. Substituting into (5), only vectors $\vec{m}$ with all coordinates weakly greater than $-A$ contribute, and grouping terms according to $s = m_1$ (permissible by easy comparability),

$$\sum_{t=-\infty}^{\infty} N_t \left(U_t(\mathbf{u}) - U_t(\tilde{\mathbf{u}})\right) = \sum_{s=-A}^{\infty} N_s \Delta(s) \left(\sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k;\, \vec{m} \geq -A,\, m_1 = s} \varphi_{\tau,\vec{m}} \right)$$

$$= \sum_{s=-A}^{\infty} N_s \Delta(s) b_s = b_0 \sum_{s=-A}^{\infty} \left(\frac{b_s}{b_0}\right) N_s \Delta(s)$$

(as in the proof of Lemma 1, the inner sums are well-defined whenever $\Delta(s) \neq 0$, by easy comparability, and the corresponding terms vanish whenever $\Delta(s) = 0$). Because $0 < b_0 < \infty$, the left-hand side is strictly positive if and only if $\sum_{s=-A}^{\infty} (b_s/b_0) N_s \Delta(s) > 0$, and the result follows from (4). □

B Total Utilitarianism

Let $\mathbf{N} = \{N_t\}_{t=-\infty}^{\infty}$ denote a population stream. Consider an extension in which all-things-considered utilities depend both on $\mathbf{N}$ and on $\mathbf{u}$, as in

$$U_t(\mathbf{N}, \mathbf{u}) = \sum_{k=0}^{\infty} \sum_{\vec{m} \in \mathbb{Z}^k} \varphi_{t,\vec{m}}(\mathbf{N}) \left(\frac{N_{m_1}}{N_t}\right) u(\vec{m}), \text{ where } \varphi_{t,t} \equiv 1$$

Here, the population trajectory is no longer fixed, so the dependence of $\varphi_{t,\vec{m}}$ on $\mathbf{N}$ is non-degenerate. For example, in the Barro and Becker model, the parameter $\varphi_{t,t+n}$ will depend on the fertility rate at time t (the ratio N_{t+n}/N_t).

Theorem 2 generalizes to variable population models, in which the social welfare function is defined over elements of the form $(\mathbf{N}, \mathbf{u})$, and is given by

$$W(\mathbf{N}, \mathbf{u}) = \sum_{t=-\infty}^{\infty} N_t U_t(\mathbf{N}, \mathbf{u}) \tag{6}$$

Where the representation applies to all pairs $(\mathbf{N}, \mathbf{u})$ for which the infinite sum converges.

The generalization is as follows.

Theorem 4. *Fix some* $\mathbf{N}$ *and* $\mathbf{u}$*, and assume that* $\mathbf{u}$ *is transparent, and that*

$$\sum_{t=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k}\varphi_{t,\vec{m}}(\mathbf{N})N_{m_1}\left|u(\vec{m})\right|<\infty \tag{7}$$

Define

$$a_t(\mathbf{N}):=\sum_{\tau=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k;\,m_1=t}\varphi_{\tau,\vec{m}}(\mathbf{N})$$

and assume that $0<a_t(\mathbf{N})<\infty$*. Then,*

$$W(\mathbf{N},\mathbf{u})=\sum_{t=-\infty}^{\infty}a_t(\mathbf{N})N_tu(t)$$

Proof. The proof follows the same steps as the proof of Lemma 1 (replacing $\Delta(s)$ with $u(s)$). Because $\mathbf{u}$ is transparent, $u(\vec{m})=u(m_1)$ for every $\vec{m}$. Substituting into (6) and grouping terms according to $s=m_1$ (which is permissible because of (7)),

$$\sum_{t=-\infty}^{\infty}N_tU_t(\mathbf{N},\mathbf{u})=\sum_{s=-\infty}^{\infty}N_su(s)\left(\sum_{\tau=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k;\,m_1=s}\varphi_{\tau,\vec{m}}(\mathbf{N})\right)=\sum_{s=-\infty}^{\infty}a_s(\mathbf{N})N_su(s)$$

□

The analogue of Theorem 1 is immediate: assume that there exist some $\{\psi_{\vec{m}}\}$ such that, for all t and $\vec{m}$, it holds that $\varphi_{t,t+\vec{m}}(\mathbf{N})=\psi_{\vec{m}}$. This means that the altruism parameters are time and population invariant: the degree to which I care about the consumption utility of my child is the same as the degree to which my child cares about his own child's consumption utility, regardless of how many children each of us has. In this case, $a_t(\mathbf{N})$ is independent both from t and from $\mathbf{N}$, and thus there exists some $a>0$ such that $a_t(\mathbf{N})=a$ for all t and $\mathbf{N}$. The above result then yields the equality

$$W(\mathbf{N},\mathbf{u})=a\sum_{t=-\infty}^{\infty}N_tu(t)$$

Hence, in this case, the total utilitarian criterion is represented both by the sum of all-things-considered utilities ($\sum_t N_tU_t$) and the sum of consumption utilities ($\sum_t N_tu(t)$).

Similar steps also yield a formal analogue of Theorem 3 (again, following the same steps as the proof and replacing $\Delta(s)$ with $u(s)$). However, this generalization is less straightforward conceptually in cases in which the dependence of $\varphi_{t,\vec{m}}$ on $\mathbf{N}$ is non-degenerate. For deceased people, it makes sense to treat both their beliefs about $\mathbf{u}$ and their beliefs about $\mathbf{N}$ as immutable. This requires a richer framework, in which the parameters $\{\varphi_{t,\vec{m}}\}$ may depend not only on the realized population trajectory, but also on each generation's beliefs and higher-order beliefs about the population trajectory. Such an extension is possible, but beyond the scope of this paper.

C Egalitarian Considerations

While this paper is focused on the classical utilitarian criterion, it also has interesting implications for variants of utilitarianism that incorporate egalitarian concerns. In this section, I consider the rank-discounted utilitarian approach in Zuber and Asheim [2012]. This normative criterion assigns weights to different generations' utilities in a way that is decreasing in their living standards — expressing the "prioritarian" ethical stance that we ought to prioritize improvements in the well-being of worse off individuals.

For the purpose of this extension, it is useful to restrict attention to outcome-and-belief profiles in which both $u(t)$ and U_t are increasing in t (in line with the observation that living standards have been increasing over time). The rank of generation t is therefore given simply by t (where ranks go from $-\infty$ to ∞). Assume that the (restricted) social ordering of outcome-and-belief profiles in which $u(t)$ and U_t increasing in t is given by

$$\sum_{t=-\infty}^{\infty}\left(\frac{1}{1+\rho}\right)^{t} N_t(U_t(\mathbf{u})-U_t(\tilde{\mathbf{u}}))>0 \Leftrightarrow \mathbf{u}\succ\tilde{\mathbf{u}}$$

where $\rho>0$. This is a rank-discounted utilitarian criterion, in which better-off generations receive a lower weight.

Note that

$$\sum_{t=-\infty}^{\infty}\left(\frac{1}{1+\rho}\right)^{t} N_t(U_t(\mathbf{u})-U_t(\tilde{\mathbf{u}})) = \sum_{t=-\infty}^{\infty} N_t\left(\left(\left(\frac{1}{1+\rho}\right)^{t} U_t(\mathbf{u})\right)-\left(\left(\frac{1}{1+\rho}\right)^{t} U_t(\tilde{\mathbf{u}})\right)\right)$$

$$= \sum_{t=-\infty}^{\infty} N_t\left(\left(\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k}\tilde{\varphi}_{t,\vec{m}}\left(\frac{N_{m_1}}{N_t}\right)u(\vec{m})\right)-\left(\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k}\tilde{\varphi}_{t,\vec{m}}\left(\frac{N_{m_1}}{N_t}\right)\tilde{u}(\vec{m})\right)\right)$$

where

$$\tilde{\varphi}_{t,\vec{m}} := \left(\frac{1}{1+\rho}\right)^{t}\varphi_{t,\vec{m}}.$$

This model is therefore formally equivalent to the utilitarian model, with the modification that $\varphi_{t,\vec{m}}$ is replaced by $\tilde{\varphi}_{t,\vec{m}}$.

The easy-comparability condition is modified as follows.

Definition. *Two outcome-and-belief profiles* $\mathbf{u}$ *and* $\tilde{\mathbf{u}}$ *are* easily comparable in the rank-discounted model *if*

$$\sum_{t=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k}\tilde{\varphi}_{t,\vec{m}}N_{m_1}\left|u(\vec{m})-\tilde{u}(\vec{m})\right|<\infty$$

Defining $\tilde{a}_t$ as the analogue of a_t, we have that

$$\tilde{a}_t := \sum_{\tau=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t}\tilde{\varphi}_{\tau,\vec{m}} = \sum_{\tau=-\infty}^{\infty}\sum_{k=1}^{\infty}\sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t}\left(\frac{1}{1+\rho}\right)^{\tau}\varphi_{\tau,\vec{m}}$$

$$= \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=t} \left(\frac{1}{1+\rho}\right)^{t-\tau} \varphi_{t-\tau,\vec{m}} = \left(\frac{1}{1+\rho}\right)^{t} \sum_{\tau=-\infty}^{\infty} \sum_{k=1}^{\infty} \sum_{\vec{m}\in\mathbb{Z}^k;\, m_1=\tau} \left(\frac{1}{1+\rho}\right)^{-\tau} \varphi_{t-\tau,t-\tau+\vec{m}}$$

In particular, when intergenerational altruism is time invariant ($\varphi_{t,t+\vec{m}} = \varphi_{t',t'+\vec{m}}$ for any $\vec{m}$), the constants $\tilde{a}_t$ are proportional to $(1/(1+\rho))^t$.

By Theorem 2, if $0 < \tilde{a}_t < \infty$ for all t and $\mathbf{u}$ and $\tilde{\mathbf{u}}$ are two transparent outcome-and-belief profiles that are easily comparable in the rank-discounted model, then

$$\sum_{t=-\infty}^{\infty} \left(\frac{\tilde{a}_t}{\tilde{a}_0}\right) N_t\left(u(t) - \tilde{u}(t)\right) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}}$$

Thus, when altruism is time-invariant, the constants $\tilde{a}_t$ are proportional to $(1/(1+\rho))^t$ and hence we obtain the discounted-utilitarian condition

$$\sum_{t=-\infty}^{\infty} \left(\frac{1}{1+\rho}\right)^{t} N_t\left(u(t) - \tilde{u}(t)\right) > 0 \Leftrightarrow \mathbf{u} \succ \tilde{\mathbf{u}}$$

We thus obtain a clean analogue of Theorem 1: rank-discounted utilitarianism in U_t is equivalent to rank-discounted utilitarianism in $u(t)$. This illustrates that egalitarian considerations continue to provide reasons to discount the consumption utility of future people, even in the presence of intergenerational altruism.