Available online at www.sciencedirect.com

SciVerse ScienceDirect

Journal of Economic Theory 147 (2012) 1572–1601

JOURNAL OF Economic Theory

www.elsevier.com/locate/jet

Justifying social discounting: The rank-discounted utilitarian approach ☆

Stéphane Zuber [a,b,*], Geir B. Asheim [c]

[a] CERSES, Université Paris Descartes and CNRS, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France
[b] CESifo, Germany
[c] Department of Economics, University of Oslo, Norway

Received 19 March 2010; final version received 24 June 2011; accepted 24 August 2011

Available online 19 September 2011

Abstract

The discounted utilitarian criterion for infinite horizon social choice has been criticized for treating generations unequally. We propose an extended rank-discounted utilitarian (ERDU) criterion instead. The criterion amounts to discounted utilitarianism on non-decreasing streams, but it treats all generations impartially: discounting becomes the mere expression of intergenerational inequality aversion. We show that more inequality averse ERDU societies have higher social discount rates when future generations are better off. We apply the ERDU approach in two benchmark economic growth models and prove that it promotes sustainable policies that maximize discounted utilitarian welfare.

JEL classification: D63; H43; Q56

Keywords: Intergenerational equity; Social discounting; Discounted utilitarianism; Sustainability

☆ We thank Antoine Bommier, Wolfgang Buchholz, Marc Fleurbaey, Christian Gollier, Luc Lauwers, Michel Le Breton, François Maniquet, John Roemer, an associate editor, two anonymous referees and seminar audiences in Berlin, Louvain, Montréal, Moscow, Oslo, Paris and Toulouse for their comments and useful references. Zuber warmly thanks CORE, Université catholique de Louvain, where he was a postdoctoral research fellow at the time this paper was written and he acknowledges the financial support of the Chair Lhoist Berghmans in Environmental Economics. This paper is part of the research activities at the Centre for the Study of Equality, Social Organization, and Performance (ESOP) at the Department of Economics at the University of Oslo. ESOP is supported by the Research Council of Norway.

* Corresponding author at: CERSES, Université Paris Descartes and CNRS, 45 rue des Saints-Pères, F-75270 Paris Cedex 06, France. Fax: +33 (0)1 42 86 42 41.

E-mail addresses: stephane.zuber@parisdescartes.fr (S. Zuber), g.b.asheim@econ.uio.no (G.B. Asheim).

0022-0531/$ – see front matter
doi:10.1016/j.jet.2011.08.001

1. Introduction

The most popular objective function used to determine optimal policies in infinite horizon models is the discounted utilitarian criterion,

$$\sum_{t\in\mathbb{N}} \beta^{t-1} u(x_t), \tag{1}$$

where $0 < \beta < 1$ is the utility discount factor and x_t is the consumption of generation t. This criterion has been heavily criticized on the ground that it treats successive generations differently. Many economists in the utilitarian tradition have denounced this deviation from the ideal of equal treatment of all individuals. For instance, Frank Ramsey famously described discounting as a "practice which is ethically indefensible and arises merely from the weakness of the imagination" [37, p. 543]. Among others, Pigou [35] and Harrod [29] have also stigmatized utility discounting.

Drawing on these criticisms, a prolific literature has studied whether it would be possible to combine the principle of procedural equity (equal treatment of all generations) with the Pareto principle in the context of infinite consumption streams. Although some positive results have been obtained, most of this literature stemming from Diamond [20] has reached negative conclusions [8,51,31].

At the same time, several authors have pointed out the distributional consequences of not discounting future generations' utility. Mirrlees [32] computed optimal intertemporal consumption patterns in plausible economic models using the undiscounted utilitarian criterion (the so-called Ramsey criterion). He observed that present generations should save up to 50% of their net income for the sake of future generations. The finding was best summarized by philosopher John Rawls who declared that "the utilitarian doctrine may direct us to demand heavy sacrifices of the poorer generations for the sake of greater advantages for the later ones that are far better off" [38, p. 253]. He went on to say that "these consequences can be to some degree corrected by discounting the welfare of those living in the future" [38, p. 262].

Although Rawls did not endorse discounted utilitarianism (for the very reason that its failure to comply with procedural equity "has no intrinsic ethical appeal" [38, p. 262]), most of the economic literature has adopted it as the lesser of two evils. Yet the conflict between procedural equity and distributional equity in a utilitarian context has remained unsolved.

The above distributional justification for discounted utilitarianism critically relies on the assumption that future generations are better off in the implemented intergenerational allocation. However, as demonstrated by Dasgupta and Heal [18, Chapter 10], in certain technological contexts like the Dasgupta–Heal–Solow model of capital accumulation and resource depletion, discounted utilitarianism implies that generations in the distant future will be *worse* off than the present (see also [3, Section V]). Undiscounted utilitarianism may then yield more satisfactory recommendations than discounted utilitarianism. The key point is that, for utility discounting to prevent high sacrifices for the sake of others that are better off, it is critical that generations' position in time corresponds to their rank in well-being.

If we retain the interpretation of the utility discount factor as preventing high sacrifices from the poor, it looks closely related to the social weights used in rank-dependent measures of social welfare. An example of a rank-dependent criterion is the Gini social welfare function. Generalizations thereof have been proposed by Weymark [49] and Ebert [23]. The main feature of rank-dependent social welfare functions is that they put more weight on the utility of the worse off. Rank-dependent weights simply represent the society's aversion to inequality.

In this paper, we propose to apply rank-dependent methods to intergenerational justice.[1] More precisely, we put forward the proposal that the social observer use an element from the class of rank-discounted utilitarian social welfare functions:

$$\sum_{r\in\mathbb{N}} \beta^{r-1} u(x_{[r]}).$$

Here, the consumption stream $(x_{[1]}, x_{[2]}, \ldots, x_{[r]}, \ldots)$ is a reordering of the consumption stream $(x_1, x_2, \ldots, x_t, \ldots)$ such that $x_{[1]} \leqslant x_{[2]} \leqslant \cdots \leqslant x_{[r]} \leqslant \cdots$. The scalar β is a *rank* utility discount factor rather than a *time* utility discount factor and therefore does not entail partiality in favor of current generations.

However, an obstacle to applying rank-discounted utilitarianism in the context of infinite consumption streams is that some consumption streams cannot be reordered into a non-decreasing stream. The stream

$$(1, 0, 0, 0, \ldots, 0, \ldots)$$

is one example, where the first location with consumption 1 will end up at some finite location in any reordered stream. We resolve this problem by showing how the rank-discounted utilitarian approach can be extended in a natural manner to the full domain by including also consumption streams that cannot be reordered into non-decreasing streams.

This extended rank-discounted utilitarian approach coincides with discounted utilitarianism on the set of non-decreasing consumption streams. Utility discounting is then justified as an expression of inequality aversion when future generations are better off. However, and contrary to the discounted utilitarian approach, extended rank-discounted utilitarianism also satisfies procedural equity: two intergenerational consumption streams that are identical up to a permutation are deemed equally good. Furthermore, it satisfies the Strong Pareto principle on the domain of streams that can be reordered into non-decreasing streams. Hence, the extended rank-discounted utilitarian approach overcomes the impossibility results in the tradition of Diamond [20] on this domain.

In Section 3, we offer a complete characterization of extended rank-discounted utilitarian preferences. This characterization is clearly related to Koopmans' [30] characterization of discounted utilitarian preferences. The difference is that his separability and stationarity axioms are imposed on non-decreasing streams only. Separability axioms on ordered streams are common in the theory of decision under risk [36,50], in the theory of decision under uncertainty [25,41,46] and in the theory of inequality measurement [49,23]. With the exception of Rébillé [39], they have never been used in the theory of intertemporal decision making yet. They permit utilities to be weighted according to their rank in a distribution, which is exactly what rank-discounted utilitarian criteria do.

In Section 4, we provide conditions for a social observer using an extended rank-discounted utilitarian criterion to be inequality averse, in the sense that she always prefers a consumption stream obtained from another through a Pigou–Dalton redistributive transfer. We also provide conditions for comparing two social observers in terms of inequality aversion. When the social observer has homothetic preferences, these conditions are very simple: she needs to discount ranks more and to use a more concave utility function.

[1] Alternatively, the analysis could have been motivated as an extension of known results on finite rank-dependent social evaluation to the infinite case. For a comparison of our analysis with relevant contributions in this alternative setting, see Section 3.2.

Distributional equity in the spirit of Atkinson [7] has been addressed in many papers in the literature on intergenerational equity (see, e.g., [9,12,28]). However, this literature did not emphasize the effects of inequality aversion on society's choice. We claim that inequality aversion is a central notion for evaluating intergenerational problems.

In Section 5, we explore the implications of rank-discounted utilitarian social welfare functions for the social discount rate. The highly publicized debates on the social discount rate in the context of climate change have highlighted its importance for policy evaluation. An 'ethical' view has suggested low values for the social discount rate, on the ground that time utility discounting violates procedural equity. Rank-discounted utilitarianism suggests an alternative 'ethical' view where rank utility discounting is an expression of society's aversion to inequality.

Indeed, we prove that a more inequality averse social observer always discount the future more, *provided that future generations are better off*. This has important policy implications. If future generations are expected to be better off in spite of climate change, then a more inequality averse extended rank-discounted utilitarian social observer will agree with the recommendation of Nordhaus [34] to have gradual emission control policies rather than that of Stern [43] who calls for immediate action. However, since rank-discounting depends on a generation's rank in the intergenerational distribution rather than its position in time, if future generations are expected to be less well-off because of climate change, then the social discount rate should on the contrary be negative, and strong action should be undertaken to mitigate climate change.

In Section 6, we show that the extended rank-discounted utilitarian approach can be applied to find the optimal growth policy in two benchmark models: the Ramsey growth model and the Dasgupta–Heal–Solow model of capital accumulation and resource depletion. Also in these applications, inequality aversion plays a crucial role. Indeed, in a more inequality averse society, growth is prevented for a greater set of initial conditions: if the initial stock of capital is high enough, the society prefers to maintain consumption forever. Then more inequality aversion yields greater equality and lower long-run consumption.

To reach these conclusions, we start in Section 2 by introducing the framework of our analysis.

2. The framework

Let $\mathbb{N}$ denote as usual the set of natural numbers $\{1, 2, 3, \ldots\}$. Let $\mathbb{R}$ denote the set of real numbers, $\mathbb{R}_+$ the set of nonnegative real numbers, and $\mathbb{R}_{++}$ the set of positive real numbers.

Denote by $\mathbf{x} = (x_1, x_2, \ldots, x_t, \ldots)$ an infinite stream (or allocation), where $x_t \in \mathbb{R}_+$ is a one-dimensional indicator of the well-being of generation t. We refer to this indicator as the consumption of generation t, restrict attention to allocations consisting of bounded consumption streams, and denote by

$$\mathbf{X} = \left\{ \mathbf{x} = (x_1, \ldots, x_t, \ldots) \in \mathbb{R}_+^{\mathbb{N}}\colon \sup_t x_t < +\infty \right\}$$

the set of possible allocations.

For $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, write $\mathbf{x} \geqslant \mathbf{y}$ whenever $x_t \geqslant y_t$ for all $t \in \mathbb{N}$; write $\mathbf{x} > \mathbf{y}$ if $\mathbf{x} \geqslant \mathbf{y}$ and $\mathbf{x} \neq \mathbf{y}$; and write $\mathbf{x} \gg \mathbf{y}$ whenever $x_t > y_t$ for all $t \in \mathbb{N}$. For any $T \in \mathbb{N}$ and $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, denote by $\mathbf{x_T y}$ the consumption stream $\mathbf{z}$ such that $z_t = x_t$ for all $t \leqslant T$ and $z_t = y_t$ for all $t > T$. For any $x \in \mathbb{R}_+$ and $\mathbf{y} \in \mathbf{X}$, denote by $(x, \mathbf{y})$ the stream $(x, y_1, y_2, \ldots)$.

Three subsets of $\mathbf{X}$ will be of particular interest. First, we introduce the set of stationary consumption streams, $\mathbf{X^c} = \{\mathbf{x^c},\ x \in \mathbb{R}_+\}$, where for any $x \in \mathbb{R}_+$, $\mathbf{x^c} \in \mathbf{X}$ denotes the allocation such that $x_t^c = x$ for all $t \in \mathbb{N}$.

A second subset of $\mathbf{X}$ is the set of non-decreasing streams in $\mathbf{X}$. This set is denoted by $\mathbf{X}^+ = \{\mathbf{x} \in \mathbf{X}\colon x_t \leqslant x_{t+1},\ \forall t \in \mathbb{N}\}$.

The third subset of $\mathbf{X}$, playing a key role in the remainder of the paper, is the set of allocations, $\overline{\mathbf{X}}$, whose elements can be permuted to obtain non-decreasing streams. To introduce $\overline{\mathbf{X}}$ formally, let Π be the set of all permutations on $\mathbb{N}$. For any $\pi \in \Pi$ and $\mathbf{x} \in \mathbf{X}$, let $\mathbf{x}_\pi = (x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(t)}, \ldots)$. The set $\overline{\mathbf{X}}$ is defined as follows: $\overline{\mathbf{X}} = \{\mathbf{x} \in \mathbf{X}\colon \exists \pi \in \Pi,\ \mathbf{x}_\pi \in \mathbf{X}^+\}$.

The following inclusions hold: $\mathbf{X}^{\mathbf{c}} \subset \mathbf{X}^+ \subset \overline{\mathbf{X}} \subset \mathbf{X}$. In a finite setting, $\overline{\mathbf{X}}$ would be the same as $\mathbf{X}$. In the introduction we have already used the stream $(1, 0, 0, 0, \ldots, 0, \ldots)$ to illustrate why the inclusion is strict in an infinite setting.

To characterize the set $\overline{\mathbf{X}}$, let $\ell(\mathbf{x})$ denote $\liminf_{t\to+\infty} x_t$ for any $\mathbf{x} \in \mathbf{X}$. Because streams in $\mathbf{X}$ are bounded, $\ell(\mathbf{x})$ is well defined for all $\mathbf{x} \in \mathbf{X}$. Write $L(\mathbf{x}) = \{t \in \mathbb{N}\colon x_t < \ell(\mathbf{x})\}$ and denote by $|L(\mathbf{x})|$ the cardinality of $L(\mathbf{x})$.

Proposition 1.

(a) *If an allocation $\mathbf{x} \in \mathbf{X}$ satisfies $|L(\mathbf{x})| < +\infty$, then $\mathbf{x}$ belongs to $\overline{\mathbf{X}}$ if and only if $x_t \leqslant \ell(\mathbf{x})$ for all $t \in \mathbb{N}$.*

(b) *If an allocation $\mathbf{x} \in \mathbf{X}$ satisfies $|L(\mathbf{x})| = +\infty$, then $\mathbf{x}$ belongs to $\overline{\mathbf{X}}$ if and only if $x_t < \ell(\mathbf{x})$ for all $t \in \mathbb{N}$.*

Proposition 1 is clearly equivalent to the following lemma.

Lemma 1. *An allocation $\mathbf{x} \in \mathbf{X}$ belongs to $\overline{\mathbf{X}}$ if and only if the cardinality of $\Lambda_\tau(\mathbf{x}) = \{t \in \mathbb{N},\ t > \tau\colon x_t < x_\tau\}$ is finite for all $\tau \in \mathbb{N}$.*

Proof. If $|\Lambda_\tau(\mathbf{x})| = +\infty$ for some $\tau \in \mathbb{N}$, then, for any $\pi \in \Pi$, $\pi(\tau) < +\infty$ and it is impossible that $\pi(t) < \pi(\tau)$ for all $t \in \Lambda_\tau(\mathbf{x})$. Hence, $\mathbf{x} \notin \overline{\mathbf{X}}$.

Conversely, assume $|\Lambda_\tau(\mathbf{x})| < +\infty$ for all $\tau \in \mathbb{N}$. The set $\Lambda_1(\mathbf{x})$ is finite and can be reordered in non-decreasing order. These coordinate will form the n_1 first elements of the ordered stream, with $n_1 = |\Lambda_1(\mathbf{x})|$. And $\pi(1) = n_1 + 1$. Then let τ_2 be the first period such that $x_{\tau_2} \geqslant x_1$. The set $\Lambda_{\tau_2}(\mathbf{x}) \setminus \Lambda_1(\mathbf{x})$ is finite and can be ordered in non-decreasing order. These will form the n_2 next elements in the ordered stream, with $n_2 = |\Lambda_{\tau_2}(\mathbf{x})| - |\Lambda_1(\mathbf{x})|$. And $\pi(t_2) = n_1 + n_2 + 2$. Pursuing this procedure leads to an ordered stream. Hence, $\mathbf{x} \in \overline{\mathbf{X}}$. □

For $\mathbf{x} \in \overline{\mathbf{X}}$, denote by $\mathbf{x}_{[\,]} = (x_{[1]}, x_{[2]}, \ldots, x_{[r]}, \ldots)$ the non-decreasing allocation which is a permutation of $\mathbf{x}$; i.e., for some $\pi \in \Pi$ such that $\mathbf{x}_\pi \in \mathbf{X}_+$, it holds that $x_{[r]} = x_{\pi(r)}$ for all $r \in \mathbb{N}$. Note that the permutation π need not be unique (if, for instance, $x_t = x_{t'}$ for some $t \neq t'$), but the resulting non-decreasing allocation $\mathbf{x}_{[\,]}$ is unique. Likewise, for $\mathbf{x} \in \mathbf{X}$, denote by $(x_{[1]}, \ldots, x_{[|L(\mathbf{x})|]})$ the non-decreasing allocation which is a permutation of the elements of $\mathbf{x}$ satisfying $t \in L(\mathbf{x})$. The following notation is useful: $r_\tau(\mathbf{x}) = |\{t \in \mathbb{N}\colon x_t < x_\tau\}| + 1$ and $\bar{r}_\tau(\mathbf{x}) = |\{t \in \mathbb{N}\colon x_t \leqslant x_\tau\}|$. Whenever $r_\tau(\mathbf{x}) = \bar{r}_\tau(\mathbf{x}) < +\infty$, $r_\tau(\mathbf{x})$ is the unique rank of generation τ in the distribution $\mathbf{x}$. Note that whenever $r_\tau(\mathbf{x}) < +\infty$ ($\bar{r}_\tau(\mathbf{x}) < +\infty$), we have that $x_{[r_\tau(\mathbf{x})]} = x_\tau$ (and also $x_{[\bar{r}_\tau(\mathbf{x})]} = x_\tau$).

A social welfare relation (SWR) on a set $\mathbf{X}$ is a binary relation $\succsim$, where for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, $\mathbf{x} \succsim \mathbf{y}$ implies that the consumption stream $\mathbf{x}$ is deemed socially at least as good as $\mathbf{y}$. Let $\sim$ and $\succ$ denote the symmetric and asymmetric parts of $\succsim$. A social welfare function (SWF) representing

$\succsim$ is a mapping $W : \mathbf{X} \to \mathbb{R}$ with the property that for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, $W(\mathbf{x}) \geqslant W(\mathbf{y})$ if and only if $\mathbf{x} \succsim \mathbf{y}$.

3. Axiomatic foundation

The difficulty of combining equal treatment of an infinite number of generations with sensitivity to the interest of each of these generations has been the topic of a prolific literature since the seminal contribution by Diamond [20]. Although complete social preferences over infinite streams that combine equal treatment with Paretian sensitivity exist [44], they cannot be represented [8] nor explicitly described [51,31].

In this section we show how the set of ordered streams serves to overcome this impossibility. In Section 3.1, we first impose axioms sufficient to ensure numerical representability. Then we impose Paretian, separability and stationarity axioms, as used to characterize discounted utilitarianism [30], but restricted to the set of non-decreasing streams. In Section 3.2, we show how this allows us to invoke a strong axiom of equal treatment, requiring social indifference not only for finite permutations (as considered in the literature in the wake of [20]), but also for infinite permutations. In the concluding Section 3.3, we show that we are still able (i) to retain sensitivity to the interest of any one generation as long as there is only a finite number of other generations with lower consumption levels, and (ii) to satisfy other ethical axioms proposed in the literature to protect the interests of future generations.

3.1. Axioms

We first consider axioms sufficient to ensure numerical representability.

Order. The relation $\succsim$ is complete, reflexive and transitive on $\mathbf{X}$.

An SWR satisfying Order is called a social welfare order (SWO).

Continuity. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, if a sequence $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^k, \dots$ of allocations in $\mathbf{X}$ is such that $\lim_{k\to\infty} \sup_{t\in\mathbb{N}} |x_t^k - x_t| = 0$ and, for all $k \in \mathbb{N}$, $\mathbf{x}^k \succsim \mathbf{y}$ (resp. $\mathbf{x}^k \precsim \mathbf{y}$), then $\mathbf{x} \succsim \mathbf{y}$ (resp. $\mathbf{x} \precsim \mathbf{y}$).

Monotonicity. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, if $\mathbf{x} > \mathbf{y}$, then $\mathbf{x} \succsim \mathbf{y}$.

Monotonicity is implied by the Strong Pareto principle.

We then consider an axiom ensuring some sensitivity to the interests of the present generation.

Restricted Dominance. For any $x, y \in \mathbb{R}_+$, if $x > y$, then $(x, \mathbf{x^c}) \succ (y, \mathbf{x^c})$.

Restricted Dominance is implied by the Strong Pareto principle restricted to the set of streams that can be reordered into non-decreasing streams:

Restricted Strong Pareto. For any $\mathbf{x}, \mathbf{y} \in \overline{\mathbf{X}}$, if $\mathbf{x} > \mathbf{y}$, then $\mathbf{x} \succ \mathbf{y}$.

We now turn to restricted versions of the separability and stationarity axioms usually invoked to characterize discounted utilitarianism.

Restricted Separable Present. For any $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in \mathbf{X}^+$ such that (i) $x_t = x'_t$ and $y_t = y'_t$ for all $t \in \{1, 2\}$ and (ii) $x_t = y_t$ and $x'_t = y'_t$ for all $t \in \mathbb{N} \setminus \{1, 2\}$, $\mathbf{x} \succsim \mathbf{y}$ if and only if $\mathbf{x}' \succsim \mathbf{y}'$.

Restricted Separable Present is Postulate 3′a in Koopmans' [30] characterization of discounted utilitarianism restricted to the set of non-decreasing streams. We suggest that such a restriction might be supported by ethical intuition. In particular, one might accept that the stream $(1, 4, 5, 5, 5, \ldots)$ is socially better than $(2, 2, 5, 5, 5, \ldots)$, while not accepting that $(1, 4, 2, 2, 2, \ldots)$ is socially better than $(2, 2, 2, 2, 2, \ldots)$. It is not obvious that we should treat the conflict between the worst-off and the second worst-off generation presented by the first comparison in the same manner as we treat the conflict between the worst-off and the best-off generation put forward by the second comparison.

Restricted Separable Present follows from the following axiom by setting $\mathcal{T} = \{1, 2\}$.

Restricted Separability. For any $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in \mathbf{X}^+$ and any $\mathcal{T} \subset \mathbb{N}$ such that (i) $x_t = x'_t$ and $y_t = y'_t$ for all $t \in \mathcal{T}$ and (ii) $x_t = y_t$ and $x'_t = y'_t$ for all $t \in \mathbb{N} \setminus \mathcal{T}$, $\mathbf{x} \succsim \mathbf{y}$ if and only if $\mathbf{x}' \succsim \mathbf{y}'$.

Restricted Separability is closely related to the comonotonic sure-thing principle that has been introduced in the theory of decision-making under uncertainty (see [25,41,46]).

Restricted Separable Future. For any $\mathbf{x}, \mathbf{y}, \mathbf{x}', \mathbf{y}' \in \mathbf{X}^+$ such that (i) $x_t = x'_t$ and $y_t = y'_t$ for all $t \in \mathbb{N} \setminus \{1\}$ and (ii) $x_1 = y_1$ and $x'_1 = y'_1$, $\mathbf{x} \succsim \mathbf{y}$ if and only if $\mathbf{x}' \succsim \mathbf{y}'$.

Restricted Separable Future is Postulate 3b in Koopmans' [30] characterization of discounted utilitarianism restricted to the set of non-decreasing streams. It follows from Restricted Separability by setting $\mathcal{T} = \{2, 3, \ldots\}$.

Restricted Stationarity. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}^+$, there exists $z \in \mathbb{R}_+$ with $z \leqslant \min(x_1, y_1)$ such that $(z, \mathbf{x}) \succsim (z, \mathbf{y})$ if and only if $\mathbf{x} \succsim \mathbf{y}$.

Restricted Stationarity is Koopmans' [30] stationarity postulate (Postulate 4) restricted to the set of non-decreasing streams. The conjunction of Restricted Separable Future and Restricted Stationarity is the restriction of Independent Future, as used by Asheim, Mitra and Tungodden [6], to the set of non-decreasing streams.

Finally, we state the strong axiom of procedural equity, requiring social indifference with respect to all permutations $\pi \in \Pi$.

Strong Anonymity. For any $\pi \in \Pi$ and $\mathbf{x} \in \mathbf{X}$, $\mathbf{x} \sim \mathbf{x}_\pi$.

3.2. Characterization

In this subsection we characterize the class of SWOs satisfying Order, Continuity, Monotonicity, Restricted Dominance, Restricted Separable Present, Restricted Separable Future, Restricted Stationarity and Strong Anonymity. As a first step, we do so within the restricted domain $\overline{\mathbf{X}}$ of allocations that can be reordered into non-decreasing streams.

Definition 1. An SWR on $\overline{\mathbf{X}}$ is a *Rank-Discounted Utilitarian SWO* (RDU SWO) if it is represented by an SWF $\overline{W} : \overline{\mathbf{X}} \to \mathbb{R}$ defined by:

$$\overline{W}(\mathbf{x}) = (1-\beta)\sum_{r\in\mathbb{N}}\beta^{r-1}u(x_{[r]}), \tag{2}$$

where $0<\beta<1$ is a real number and the function u is continuous and increasing.

Although the RDU criterion can be seen as an infinite extension of families of single-series Ginis, as axiomatized by Bossert [11], with the Gini weight of rank r set equal to β^{r-1}, our axiomatization differs from Bossert's. The recursive methods that we use are similar to his recursivity property. However, we do not need the linear homogeneity and translatability properties which are essential for his result. We rely instead on Restricted Separable Future and Restricted Stationarity which are taken from intertemporal choice theory.[2]

Proposition 2. *If an SWR $\succsim$ on $\overline{\mathbf{X}}$ satisfies Order, Continuity, Monotonicity, Restricted Dominance, Restricted Separable Present, Restricted Separable Future, Restricted Stationarity and Strong Anonymity, then it is an RDU SWO.*

Proof. See Appendix A for a simplified version of Koopmans' [30] proof, similar to the one in Bleichrodt, Rhode and Wakker [10]. The proof is applied to non-decreasing streams, requiring the use of techniques developed by Wakker [46] for additive representation of preferences on rank-ordered sets. Continuity allows us to extend from a finite number of period to an infinite number of periods the representation on non-decreasing streams. Strong Anonymity allows us to extend the representation to the whole set $\overline{\mathbf{X}}$. □

We then turn to the demonstration of the result that this class can be characterized in terms of extended RDU SWOs on the unrestricted domain $\mathbf{X}$.

Definition 2. An SWR on $\mathbf{X}$ is an *Extended Rank-Discounted Utilitarian SWO* (ERDU SWO) if it is represented by an SWF $W:\mathbf{X}\to\mathbb{R}$ defined by:

$$W(\mathbf{x}) = u\big(\ell(\mathbf{x})\big) + (1-\beta)\sum_{r=1}^{|L(\mathbf{x})|}\beta^{r-1}\big(u(x_{[r]}) - u\big(\ell(\mathbf{x})\big)\big), \tag{3}$$

where $0<\beta<1$ is a real number and the function u is continuous and increasing.

To investigate how the ERDU SWF W extends the RDU SWF $\overline{W}$, define, for any $\mathbf{x}\in\mathbf{X}$, $\bar{\mathbf{x}}$ as follows:

$$\begin{cases} \bar{x}_t = \min\{x_t, \ell(\mathbf{x})\} \text{ for all } t\in\mathbb{N} \text{ if } |L(\mathbf{x})| < +\infty, \\ \bar{\mathbf{x}} \text{ is the subsequence of } \mathbf{x} \text{ consisting of all } x_t \text{ with } t\in L(\mathbf{x}) \text{ if } |L(\mathbf{x})| = +\infty. \end{cases}$$

Proposition 1 implies that, by construction, $\bar{\mathbf{x}}$ belongs to $\overline{\mathbf{X}}$; therefore, $\bar{\mathbf{x}}_{[\,]}$ is well defined. It follows from (2) and (3) that for all $\mathbf{x}\in\mathbf{X}$,

$$W(\mathbf{x}) = \overline{W}(\bar{\mathbf{x}}). \tag{4}$$

[2] Another class of single-series Ginis is the class of single-parameter Ginis axiomatized by Donaldson and Weymark [21], whose generalization in a continuous framework is presented in Donaldson and Weymark [22]. The finite population counterparts of the RDU criterion does not satisfy the principle of population which characterizes single-parameter Ginis.

The ERDU SWF W is consistent with the idea of constant rank-dependent discounting: any generation t with $x_t > \ell(\mathbf{x})$ if $|L(\mathbf{x})| < +\infty$ or $x_t \geqslant \ell(\mathbf{x})$ if $|L(\mathbf{x})| = +\infty$ is infinitely ranked when consumption levels are ordered in a non-decreasing sequence, in the sense that there are infinitely many generations t' with $x_{t'} < x_t$. Hence, no weight is placed on their marginal consumption.

By Order, Continuity, Monotonicity and Restricted Dominance, for all $\mathbf{x} \in \mathbf{X}$ there exists a unique scalar x_e such that $\mathbf{x_e^c} \sim \mathbf{x}$. The scalar x_e is an equally distributed equivalent and it is a representation of $\succsim$. Hence, under these axioms, the SWF $\widetilde{W} : \mathbf{X} \to \mathbb{R}$ given by

$$\widetilde{W}(\mathbf{x}) = u(x_e),$$

with the function u being continuous and increasing, is well defined.

Lemma 2. *Assume that an SWR $\succsim$ satisfies Order, Continuity, Monotonicity and Restricted Dominance, and is represented on $\overline{\mathbf{X}}$ by an RDU SWF. Then the SWR $\succsim$ is represented on $\mathbf{X}$ by the SWF $\widetilde{W}$ which coincides with $\overline{W}$ on $\overline{\mathbf{X}}$.*

Proof. An RDU SWF $(1-\beta)\sum_{r\in\mathbb{N}} \beta^{r-1} u(x_{[r]})$ where $0 < \beta < 1$ and u is increasing is a representation of the SWO $\succsim$ on $\overline{\mathbf{X}}$. Since u is increasing, $\widetilde{W}$ is a representation of $\succsim$ on $\mathbf{X}$ and, by the definition of the equally distributed equivalent on $\overline{\mathbf{X}}$, it is such that $\widetilde{W}(\mathbf{x}) = (1-\beta)\sum_{r\in\mathbb{N}} \beta^{r-1} u(x_{[r]})$ for all $\mathbf{x} \in \overline{\mathbf{X}}$. □

Lemma 3. *Assume that an SWR $\succsim$ satisfies Monotonicity and Strong Anonymity, and is represented on $\overline{\mathbf{X}}$ by $\overline{W}$ and on $\mathbf{X}$ by $\widetilde{W}$. Then, for all $\mathbf{x} \in \mathbf{X}$ with $|L(\mathbf{x})| \geqslant T \geqslant 0$,*

$$\widetilde{W}(\mathbf{x}) \leqslant \overline{W}\big(x_{\pi(1)}, x_{\pi(2)}, \ldots, x_{\pi(T)}, \ell(\mathbf{x}), \ell(\mathbf{x}), \ldots\big),$$

where, $\forall t \in \{1, \ldots, T\}$, $x_{\pi(t)} < \ell(\mathbf{x})$.

Proof. Such insensitivity for $x_t > \ell(\mathbf{x})$ is shown in Appendix A. □

Lemma 4. *Assume that an SWR $\succsim$ satisfies Monotonicity and Strong Anonymity, and is represented on $\overline{\mathbf{X}}$ by $\overline{W}$ and on $\mathbf{X}$ by $\widetilde{W}$. Then, for all $\mathbf{x} \in \mathbf{X}$, $\widetilde{W}(\mathbf{x}) = \overline{W}(\bar{\mathbf{x}}) = W(\mathbf{x})$.*

Proof. This follows from Lemma 3; see Appendix A. □

Proposition 3. *Assume that an SWR $\succsim$ satisfies Order, Continuity, Monotonicity, Restricted Dominance and Strong Anonymity, and is represented on $\overline{\mathbf{X}}$ by an RDU SWF. Then the SWR $\succsim$ is represented on $\mathbf{X}$ by an ERDU SWF.*

Proof. This result follows from Lemmata 2 and 4. □

Theorem 1. *Consider an SWR $\succsim$ on $\mathbf{X}$. The following two statements are equivalent.*

(1) *$\succsim$ satisfies Order, Continuity, Monotonicity, Restricted Dominance, Restricted Separable Present, Restricted Separable Future, Restricted Stationarity and Strong Anonymity.*
(2) *$\succsim$ is an ERDU SWO.*

Proof. (1) *implies* (2). This follows from Propositions 2 and 3. (2) *implies* (1). This is easy to establish, and its proof is left to the reader. □

3.3. Properties

By combining Order, Continuity, Monotonicity and Restricted Dominance with the unrestricted versions of separability of the present and the future and stationarity — Separable Present, Separable Future and Stationarity — one obtains a characterization of discounted utilitarianism (DU), whereby all streams $\mathbf{x}$ in $\mathbf{X}$ are ranked according to the SWF (1) (cf. [6, Proposition 9]). DU does not satisfy Strong Anonymity as an axiom of procedural equity, since the permutation of consumption may change the DU social welfare. Moreover, as pointed out by Asheim, Mitra and Tungodden [6], the DU SWF does not satisfy the following distributional equity axiom, giving priority to the future in conflicts where the present is better off than the future.

Hammond Equity for the Future. For all $x, y, w, z \in \mathbb{R}_+$, if $x > y > w > z$, then $(y, \mathbf{w^c}) \succsim (x, \mathbf{z^c})$.

Finally, as pointed out by Chichilnisky [15], DU is a dictatorship of the present, which on the domain $\mathbf{X}$ can be formalized as follows:

Dictatorship of the Present. For all $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ such that $\mathbf{x} \succ \mathbf{y}$, there exist $z \in \mathbb{R}_+$ with $x_t, y_t \leqslant z$ for all $t \in \mathbb{N}$ and $T' \in \mathbb{N}$ such that, for any $\mathbf{x}', \mathbf{y}' \in [0, z]^{\mathbb{N}}$, $({}_{\mathbf{T}}\mathbf{x}, {}_{\mathbf{T+1}}\mathbf{x}') \succ ({}_{\mathbf{T}}\mathbf{y}, {}_{\mathbf{T+1}}\mathbf{y}')$ for all $T \geqslant T'$.

Hence, a setting where Order, Continuity, Monotonicity and Restricted Dominance are invoked, at least one of Separable Present, Separable Future or Stationarity must be weakened to prevent such a dictatorship.

Non-Dictatorship of the Present. Dictatorship of the Present does not hold.

Chichilnisky [15] allows for Non-Dictatorship of the Present by dropping Stationarity in the class of sustainable preferences characterized by her Theorem 2. However, SWOs in this class do not satisfy the two other ethical axioms: Hammond Equity for the Future and Strong Anonymity.

Building on Asheim, Mitra and Tungodden's [6] axiomatic analysis of sustainable recursive SWFs, Asheim and Mitra [5] allow for Hammond Equity for the Future by restricting Separable Present to the set of non-decreasing streams (i.e., imposing Restricted Separable Present) in their analysis of sustainable discounted utilitarian (SDU) SWOs, while retaining the remaining axioms of the above axiomatization of DU. Moreover, SDU SWOs satisfy Non-Dictatorship of the Present, but fail to satisfy Strong Anonymity.

With this background, it is of interest to note the following proposition.

Proposition 4. *An ERDU SWO satisfies Hammond Equity for the Future and Non-Dictatorship of the Present.*

Proof. *An ERDU SWO satisfies Hammond Equity for the Future.* Let $x > y > w > z \geqslant 0$. Then $W(y, \mathbf{w^c}) = u(w) > u(z) = W(x, \mathbf{z^c})$. *An ERDU SWO satisfies Non-Dictatorship of the Present.* Let $\mathbf{x} \succ \mathbf{y}$. Choose any $z \geqslant 0$ satisfying $x_t, y_t \leqslant z$ for all $t \in \mathbb{N}$. Let $\mathbf{x}' = \mathbf{y}' = \mathbf{0^c} \in [0, z]^{\mathbb{N}}$. Then $W({}_{\mathbf{T}}\mathbf{x}, {}_{\mathbf{T+1}}\mathbf{x}') = W({}_{\mathbf{T}}\mathbf{y}, {}_{\mathbf{T+1}}\mathbf{y}')$ for all $T \geqslant 0$. □

Hence, when moving from SDU to ERDU, Strong Anonymity is added and Separable Future and Stationarity are weakened to Restricted Separable Future and Restricted Stationarity. The weakening of Stationarity to Restricted Stationarity means that we lose time-consistency when social preferences are time-invariant. Even though time-inconsistency turns out *not* to be an issue when ERDU SWFs are applied to the Ramsey and Dasgupta–Heal–Solow growth models, as we do in Section 6, it might be a problem in other environments. It also excludes the use of recursive methods, e.g., when faced with uncertainty.

Still, it is remarkable that anonymity (even in its strongest form, Strong Anonymity, allowing infinite permutations) can be combined with numerical representability and some sensitivity to the interests of the present generation, as such attempts have not previously lead to SWOs with attractive properties.[3] Strong Anonymity is a basic form of procedural equity, corresponding to equal treatment of generations. In this sense it seems more fundamental than the distributional axiom Hammond Equity for the Future.

As pointed out by Van Liedekerke and Lauwers [45], Strong Anonymity is in conflict with the Strong Pareto principle. Moreover, Basu and Mitra [8] showed that even Finite Anonymity (i.e., anonymity in its weaker form, involving only finite permutations) rules out the Strong Pareto principle when combined with numerical representability. Finally, Zame [51] and Lauwers [31] demonstrated that SWOs satisfying both Finite Anonymity and Strong Pareto cannot be explicitly described.

Strong Anonymity is even in conflict with the Weak Pareto principle whereby one stream is preferred to another stream if the former has higher consumption than the latter at all times. This is demonstrated by the following adaptation of Fleurbaey and Michel's [24] proof of their Theorem 1 to a setting where streams are bounded. For this purpose, consider

$$\mathbf{x} = \left(\frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \dots, \frac{1}{k+2}, \frac{k+1}{k+2}, \dots\right),$$

$$\mathbf{y} = \left(\frac{1}{4}, \frac{1}{3}, \frac{1}{5}, \frac{2}{3}, \dots, \frac{1}{k+3}, \frac{k}{k+1}, \dots\right).$$

Then, by Strong Anonymity, $\mathbf{x}$ is indifferent to $\mathbf{y}$ since $\mathbf{x}$ is a permutation of $\mathbf{y}$ (move location 2 to location 1, all other even locations two periods backwards, and all odd locations two periods forwards). Still, $x_t > y_t$ for all $t \in \mathbb{N}$.

Because ERDU SWOs satisfy Strong Anonymity, it follows that they must be in conflict with the Weak Pareto principle on the full domain $\mathbf{X}$, which indeed is what Lemma 3 entails. However, an important feature of ERDU SWOs is that they satisfy the Strong Pareto principle on the restricted set $\overline{\mathbf{X}}$ of streams that can be permuted into non-decreasing streams. This means that ERDU SWOs retain sensitivity to the interest of any one generation as long as there is only a finite number of other generations with lower consumption levels. Moreover, they fulfill the separability axiom on the set $\mathbf{X}^+$ of non-decreasing streams. These are straightforward consequences of Eq. (3), so that no proof is provided.

Proposition 5. *An ERDU SWO satisfies Restricted Strong Pareto and Restricted Separability.*

[3] Sakai's [40] nice characterization of a class of welfare functions depending only on limsup and liminf is a recent contribution combining Strong Anonymity and representability. However, it is insensitive to the consumption of any finite subset of generations.

4. Inequality aversion

Up to now, we have addressed the issue of procedural equity and its compatibility with the sensitivity to the interests of each generation.

In this section, we introduce concerns for distributional equity. We will show that inequality aversion can be properly measured and compared within the ERDU class of preferences. The next two sections will then show that inequality aversion has significant policy implications.

4.1. The Pigou–Dalton Transfer principle and inequality aversion

Following the practice of expressing distributional equity ideals by means of transfer axioms, we consider a weak form of the Pigou–Dalton Transfer principle:

Pigou–Dalton Transfer. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, if there exist $\varepsilon \in \mathbb{R}_{++}$ and $\tau, \tau' \in \mathbb{N}$ such that $\varepsilon \leqslant y_\tau + \varepsilon = x_\tau \leqslant x_{\tau'} = y_{\tau'} - \varepsilon$ and $y_t = x_t$ for all $t \neq \tau, \tau'$, then $\mathbf{x} \succsim \mathbf{y}$.

In this subsection, we study the restrictions imposed by Pigou–Dalton Transfer on ERDU criteria. These restrictions hold on the rank utility discount factor β and on the utility function u in Eq. (3). Write $\succsim_{\beta,u}$ for the ERDU SWO characterized by β and u.

Introduce the following index of non-concavity of the function u (which, recall, is continuous and increasing):

$$\mathcal{C}_u = \sup_{0<\varepsilon \leqslant x \leqslant x'} \frac{u(x'+\varepsilon) - u(x')}{u(x) - u(x-\varepsilon)}.$$

As shown by Chateauneuf, Cohen and Meilijson [14], this index has two interesting properties: (1) $\mathcal{C}_u \geqslant 1$, with $\mathcal{C}_u = 1$ corresponding to u being concave; and (2) when u is differentiable, $\mathcal{C}_u = \sup_{y \leqslant x}(u'(x)/u'(y))$.

The non-concavity index $\mathcal{C}_u$ and the rank utility discount factor β jointly characterize ERDU SWFs satisfying the Pigou–Dalton Transfer principle.

Proposition 6. *An ERDU SWO $\succsim_{\beta,u}$ on $\mathbf{X}$ satisfies Pigou–Dalton Transfer if and only if*

$$\beta \times \mathcal{C}_u \leqslant 1.$$

Proof. See Appendix A. □

The condition $\beta \times \mathcal{C}_u \leqslant 1$ implies that the utility function u must not be 'too non-concave'. The concavity of u, though sufficient, is not necessary for an ERDU SWO to satisfy Pigou–Dalton Transfer.

In applications, it is convenient to consider the more specific class of homothetic ERDU SWOs, which yield clear-cuts results for comparisons of inequality aversion and for the expression of the discount rate.

Definition 3. An SWO $\succsim$ on $\mathbf{X}$ is a *Homothetic Extended Rank-Discounted Utilitarian SWO* (HERDU SWO) if it can be represented by an SWF $W : \mathbf{X} \to \mathbb{R}$ defined by:

$$W(\mathbf{x}) = \begin{cases} \frac{(\ell(\mathbf{x}))^{1-\eta}}{1-\eta} + (1-\beta)\sum_{r\in\mathbb{N}} \beta^{r-1}(\frac{x_{[r]}^{1-\eta}}{1-\eta} - \frac{(\ell(\mathbf{x}))^{1-\eta}}{1-\eta}) & \text{if } \eta \neq 1, \\ \ln \ell(\mathbf{x}) + (1-\beta)\sum_{r\in\mathbb{N}} \beta^{r-1}(\ln x_{[r]} - \ln \ell(\mathbf{x})) & \text{if } \eta = 1, \end{cases} \tag{5}$$

where $0 < \beta < 1$ is a real number.

Denote by $\succsim_{\beta,\eta}$ an HERDU SWO represented by an SWR W with rank utility discount factor β and utility function $u(x) = x^{1-\eta}/(1-\eta)$ (or $u(x) = \ln x$ if $\eta = 1$). In contrast to the general case, the (weak) concavity of u is necessary and sufficient for an HERDU SWOs to be inequality averse. For an HERDU SWO, it is indeed the case that $\mathcal{C}_u = 1$ whenever $\eta \geqslant 0$ and $\mathcal{C}_u = +\infty$ whenever $\eta < 0$. This is summarized in the following corollary:

Corollary 1. *An HERDU SWO on* $\mathbf{X}$ $\succsim_{\beta,\eta}$ *satisfies Pigou–Dalton Transfer if and only if* $\eta \geqslant 0$.

4.2. Comparative inequality aversion

Ranking different criteria according to the strength of their concerns for equality is an important prerequisite to study the policy implications of inequality aversion. The common way to do so is to define and compare the degree of inequality aversion of the underlying SWOs. The aim of this section is to perform such comparisons in the case of ERDU SWOs.

We follow the procedure proposed in the literature on risk/uncertainty aversion to make such comparisons (see Grant and Quiggin [27]). It consists in: (i) defining an inequality relation $\succ_I$; (ii) declaring an SWO $\succsim$ at least as inequality averse as an SWO $\widehat{\succsim}$ if, for any allocation $\mathbf{y}$, whenever a less unequal allocation $\mathbf{x}$ (according to $\succ_I$) is preferred to $\mathbf{y}$ according to $\widehat{\succsim}$, then $\mathbf{x}$ is also preferred to $\mathbf{y}$ according to $\succsim$.

We use a simple definition of the relation 'more unequal than' based on the notion of a 'local increase' in inequality, namely an inequality change affecting only two generations and leaving generations' ranks unchanged.

Definition 4. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, allocation $\mathbf{y}$ represents an *elementary increase in inequality* with respect to allocation $\mathbf{x}$, denoted by $\mathbf{y} \succ_I \mathbf{x}$, if there exist $\varepsilon, \varepsilon' \in \mathbb{R}_{++}$ and $\tau, \tau' \in \mathbb{N}$ such that $y_\tau + \varepsilon = x_\tau \leqslant x_{\tau'} = y_{\tau'} - \varepsilon'$, $r_\tau(\mathbf{y}) = r_\tau(\mathbf{x})$, $\bar{r}_{\tau'}(\mathbf{y}) = \bar{r}_{\tau'}(\mathbf{x})$, and $y_t = x_t$ for all $t \neq \tau, \tau'$.

The inequality relation $\succ_I$ is used to define comparative inequality aversion:

Definition 5. An SWO $\succsim$ is *at least as inequality averse as* an SWO $\widehat{\succsim}$ if, for any $\mathbf{x}$ and any $\mathbf{y} \succ_I \mathbf{x}$: (i) $\mathbf{x} \,\widehat{\succsim}\, \mathbf{y} \Rightarrow \mathbf{x} \succsim \mathbf{y}$; and (ii) $\mathbf{x} \,\widehat{\succ}\, \mathbf{y} \Rightarrow \mathbf{x} \succ \mathbf{y}$.

Consider two ERDU SWOs, $\succsim_{\beta,u}$ and $\succsim_{\hat{\beta},\hat{u}}$. To assess their relative inequality aversion, the rank discount factors β and $\hat{\beta}$ and the relative concavity of the utility functions u and $\hat{u}$ must be compared. The following two indices do so:

$$\mathcal{D}_{\beta,\hat{\beta}} = \inf_{t<t'} \frac{\beta^t/\hat{\beta}^t}{\beta^{t'}/\hat{\beta}^{t'}} = \begin{cases} \hat{\beta}/\beta & \text{if } \beta \leqslant \hat{\beta}, \\ 0 & \text{if } \beta > \hat{\beta}, \end{cases}$$

$$\mathcal{C}_{u,\hat{u}} = \sup_{0 \leqslant x_1 < x_2 \leqslant x_3 < x_4} \frac{[u(x_4) - u(x_3)]/[\hat{u}(x_4) - \hat{u}(x_3)]}{[u(x_2) - u(x_1)]/[\hat{u}(x_2) - \hat{u}(x_1)]}.$$

The index $\mathcal{D}_{\beta,\hat{\beta}}$ is an index of the relative decreasing speed of the social weights. The faster the social weights decrease, the less the society cares for better off generations. The index $\mathcal{C}_{u,\hat{u}}$ is an index of relative concavity of the utility functions u and $\hat{u}$. Furthermore, by [27], $\mathcal{C}_{u,\hat{u}} \geqslant 1$, with $\mathcal{C}_{u,\hat{u}} = 1$ corresponding to the case where u is an increasing concave transformation of $\hat{u}$. In addition, if u and $\hat{u}$ are differentiable, $\mathcal{C}_{u,\hat{u}} = \sup_{y \leqslant x}(u'(x)\hat{u}'(y))/(u'(y)\hat{u}'(x))$.

The comparative inequality aversion of two ERDU SWOs can be characterized as follows:

Proposition 7. *Consider two ERDU SWOs, $\succsim_{\beta,u}$ and $\succsim_{\hat{\beta},\hat{u}}$, on* $\mathbf{X}$. *Then $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$ if and only if*

$$\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}.$$

Proof. See Appendix A. □

By Proposition 7, $\beta \leqslant \hat{\beta}$ is a necessary condition for $\succsim_{\beta,u}$ to be at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$. A more inequality averse ERDU social observer has a lower rank utility discount factor and thus discounts more the utility of better off generations. Moreover, if $\beta = \hat{\beta}$, then u must be a concave transformation of $\hat{u}$.

Even clearer results can be obtained in the case of HERDU SWOs. Indeed, it is straightforward that, whenever $u(x) = x^{1-\eta}/(1-\eta)$ and $\hat{u}(x) = x^{1-\hat{\eta}}/(1-\hat{\eta})$, $\mathcal{C}_{u,\hat{u}} = 1$ if $\eta \geqslant \hat{\eta}$, and $\mathcal{C}_{u,\hat{u}} = +\infty$ if $\eta < \hat{\eta}$. We hence obtain the following simple conditions for comparative inequality aversion of HERDU SWOs:

Corollary 2. *Consider two HERDU SWOs, $\succsim_{\beta,\eta}$ and $\succsim_{\hat{\beta},\hat{\eta}}$, on* $\mathbf{X}$. *Then $\succsim_{\beta,\eta}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{\eta}}$ if and only if $\beta \leqslant \hat{\beta}$ and $\eta \geqslant \hat{\eta}$.*

As in the static case, inequality aversion is a key policy parameter in intertemporal problems, playing an important role in designing optimal policies. In Section 5, we describe how it affects social discounting, while in Section 6, we study optimal ERDU policies and highlight the impact of inequality aversion.

5. Rank-discounted utilitarianism and social discounting

Triggered by the Stern [43] review on the economics of climate change, the social discount rate has attracted much attention in recent years [33,48,16]. The controversy has not turned on the social welfare function used to assess different streams, as all the authors have endorsed the DU approach. The controversy has turned on the value of the parameters in the DU SWF (1). In particular, the time utility discount factor β and the elasticity of marginal utility, $\eta_u(x) = -d\ln u'(x)/d\ln x$, have a critical role in the determination of the social discount rate. However, there has been no consensus on the interpretation and the value of these key parameters.

In this section, we derive the social discount rate arising from ERDU SWOs. In doing so, we prove that the key parameters of the social discount rate have interpretations in terms of inequality aversion.

Assume that an ERDU SWO $\succsim_{\beta,u}$ has the property the function u in Eq. (3) is twice continuously differentiable. In that case, $\succsim_{\beta,u}$ is said to be a smooth ERDU SWO. Also, consider consumption streams $\mathbf{x}$ in $\mathbf{X}$ where

(i) $|L(\mathbf{x})| < +\infty$ and $x_t \neq \ell(\mathbf{x})$ for all $t \in \mathbb{N}$, or $|L(\mathbf{x})| = +\infty$,
(ii) no pair in $L(\mathbf{x})$ has the same consumption level (i.e., $x_t \neq x_\tau$ if $t, \tau \in L(\mathbf{x})$).

The set of such streams is denoted by $\mathbf{X}_{\neq}$. Any stream in $\mathbf{X}_{\neq}$ has the property that $r_t(\mathbf{x}) = \bar{r}_t(\mathbf{x}) < +\infty$ if $t \in L(\mathbf{x})$, while $r_t(\mathbf{x}) = \bar{r}_t(\mathbf{x}) = +\infty$ otherwise. An SWR W representing a smooth ERDU SWO is differentiable on $\mathbf{X}_{\neq}$ only, with $\partial W(\mathbf{x})/\partial x_t = \beta^{r_t(\mathbf{x})-1}u'(x_t) > 0$ if $t \in L(\mathbf{x})$ and $\partial W(\mathbf{x})/\partial x_t = 0$ otherwise.

The social discount rate evaluates how much an increase in marginal consumption in period t is 'worth' in terms of first period consumption. It is given by the following formal expression:[4]

Definition 6. Let W be the SWF used to evaluate policies. Then the *social discount rate* at period t for a stream $\mathbf{x}$ is:

$$\rho_t(\mathbf{x}) = \frac{\ln(\partial W/\partial x_1) - \ln(\partial W/\partial x_t)}{t-1}.$$

Consider a smooth ERDU SWO $\succsim_{\beta,u}$ and denote by $\delta = -\ln\beta$ the rank utility discount rate. Also denote by $g_t(\mathbf{x})$ the average per period growth rate between 1 and t: $g_t(\mathbf{x}) = (\ln x_t - \ln x_1)/(t-1)$. The social discount rate arising from a smooth ERDU SWO can now be approximated:

Proposition 8. *Let $\succsim_{\beta,u}$ be a smooth ERDU SWO and consider a stream $\mathbf{x} \in \mathbf{X}_{\neq}$ with $1 \in L(\mathbf{x})$. Then the social discount rate, $\rho_t(\mathbf{x})$, at period $t \in L(\mathbf{x})\backslash\{1\}$ is approximated by the RHS of the following expression*:

$$\rho_t(\mathbf{x}) \approx \frac{r_t(\mathbf{x}) - r_1(\mathbf{x})}{t-1}\delta + \eta_u(x_1)g_t(\mathbf{x}). \tag{6}$$

Proof. From the ERDU SWO it follows that

$$\begin{aligned}
\rho_t(\mathbf{x}) &= \frac{((r_1(\mathbf{x})-1)\ln\beta + \ln(u'(x_1))) - ((r_t(\mathbf{x})-1)\ln\beta + \ln(u'(x_t)))}{t-1} \\
&= \frac{r_t(\mathbf{x}) - r_1(\mathbf{x})}{t-1}\delta + \frac{\ln u'(x_1) - \ln u'(x_t)}{t-1} \\
&\approx \frac{r_t(\mathbf{x}) - r_1(\mathbf{x})}{t-1}\delta - \frac{d\ln u'(x_1)}{d\ln x_1}\cdot\frac{\ln x_t - \ln x_1}{t-1} \\
&= \frac{r_t(\mathbf{x}) - r_1(\mathbf{x})}{t-1}\delta + \eta_u(x_1)g_t(\mathbf{x}),
\end{aligned}$$

using a log-linear approximation for $u'(x)$. □

[4] To understand the expression, imagine that today the society makes a marginal investment ε whose rate of return is ρ, so that the generation born in period t can consume $e^{\rho(t-1)}\varepsilon$ more units of aggregate good. The change in social welfare through this investment is:

$$dW(\mathbf{x}) = \frac{\partial W}{\partial x_t}e^{\rho(t-1)}\varepsilon - \frac{\partial W}{\partial x_1}\varepsilon.$$

The social discount rate is the rate of return that makes the change in social welfare nil.

Approximation (6) shows that the social discount rate is rank-dependent: it depends crucially on the distance between the welfare rank of generation t and the one of the first generation. The further generation t is in the intergenerational distribution, the larger the social discount rate, and vice versa.

This remark leads to a second insight. If generation t is worse off than the first generation, the social discount rate will be negative, provided that $\eta_u(x_1) \geqslant 0$, which is always the case when u is concave. It has been pointed out in the literature using a DU approach that the social discount rate *may* be negative when future generations are sufficiently worse off (see, for instance, [16, p. 150]). With ERDU, this should *always* be the case as soon as future generations are worse off and the function u is concave.

On the set of increasing consumption streams, the familiar expression $\rho_t(\mathbf{x}) \approx \delta + \eta_u(x_1)g_t(\mathbf{x})$ is obtained. For smooth HERDU SWOs, the log-linear approximation of marginal utility is exact and the expression becomes $\rho_t(\mathbf{x}) = \delta + \eta g_t(\mathbf{x})$. This expression emphasizes the crucial role played by the ethical parameters to determine the social discount rate. Indeed, δ and η jointly characterize the attitude towards inequality: a more inequality averse social observer should have a higher δ (lower β) and/or a higher η. Therefore, a more inequality averse society should discount the future more whenever future generations are better off. This insight actually generalizes to all ERDU SWOs.

Proposition 9. *Consider two smooth ERDU SWOs, $\succsim_{\beta,u}$ and $\succsim_{\hat{\beta},\hat{u}}$, and a stream $\mathbf{x} \in \mathbf{X}_{\neq}$ with $1 \in L(\mathbf{x})$. Let $\rho_t(\mathbf{x})$ and $\hat{\rho}_t(\mathbf{x})$ be the associated discount rates at period $t \in L(\mathbf{x})\backslash\{1\}$. If $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$, then*:

(1) $\rho_t(\mathbf{x}) \geqslant \hat{\rho}_t(\mathbf{x})$ *if* $x_t > x_1$.
(2) $\rho_t(\mathbf{x}) \leqslant \hat{\rho}_t(\mathbf{x})$ *if* $x_t < x_1$.

Proof. For any $\mathbf{x} \in \mathbf{X}_{\neq}$ and any $t \in L(\mathbf{x})\backslash\{1\}$, $\rho_t(\mathbf{x}) \geqslant \hat{\rho}_t(\mathbf{x})$ if and only if

$$\Delta = \frac{\partial W/\partial x_1}{\partial W/\partial x_t} - \frac{\partial \widehat{W}/\partial x_1}{\partial \widehat{W}/\partial x_t} = \frac{\beta^{r_1(\mathbf{x})-1}u'(x_1)}{\beta^{r_t(\mathbf{x})-1}u'(x_t)} - \frac{\hat{\beta}^{r_1(\mathbf{x})-1}\hat{u}'(x_1)}{\hat{\beta}^{r_t(\mathbf{x})-1}\hat{u}'(x_t)} \geqslant 0,$$

where W ($\widehat{W}$) represents $\succsim_{\beta,u}$ ($\succsim_{\hat{\beta},\hat{u}}$). There are two ways to rearrange Δ:

$$\Delta = \left(\frac{\beta^{r_1(\mathbf{x})}/\hat{\beta}^{r_1(\mathbf{x})}}{\beta^{r_t(\mathbf{x})}/\hat{\beta}^{r_t(\mathbf{x})}} - \frac{u'(x_t)\hat{u}'(x_1)}{u'(x_1)\hat{u}'(x_t)}\right)\frac{u'(x_1)/u'(x_t)}{\hat{\beta}^{r_t(\mathbf{x})}/\hat{\beta}^{r_1(\mathbf{x})}}, \tag{7}$$

$$\Delta = \left(\frac{u'(x_1)\hat{u}'(x_t)}{u'(x_t)\hat{u}'(x_1)} - \frac{\beta^{r_t(\mathbf{x})}/\hat{\beta}^{r_t(\mathbf{x})}}{\beta^{r_1(\mathbf{x})}/\hat{\beta}^{r_1(\mathbf{x})}}\right)\frac{\hat{u}'(x_1)/\hat{u}'(x_t)}{\beta^{r_t(\mathbf{x})}/\beta^{r_1(\mathbf{x})}}. \tag{8}$$

Using the definitions of $\mathcal{D}_{\beta,\hat{\beta}}$ and $\mathcal{C}_{u,\hat{u}}$, we obtain: by Eq. (7), for $x_t > x_1$,

$$\Delta \geqslant (\mathcal{D}_{\beta,\hat{\beta}} - \mathcal{C}_{u,\hat{u}})\big(u'(x_1)/u'(x_t)\big)/\big(\hat{\beta}^{r_t(\mathbf{x})}/\hat{\beta}^{r_1(\mathbf{x})}\big) \geqslant 0$$

by Proposition 7, noting that $r_t(\mathbf{x}) > r_1(\mathbf{x})$; by Eq. (8), for $x_t < x_1$,

$$\Delta \leqslant (\mathcal{C}_{u,\hat{u}} - \mathcal{D}_{\beta,\hat{\beta}})\big(\hat{u}'(x_1)/\hat{u}'(x_t)\big)\big(\beta^{r_t(\mathbf{x})}/\beta^{r_1(\mathbf{x})}\big) \leqslant 0$$

by Proposition 7, noting that $r_t(\mathbf{x}) < r_1(\mathbf{x})$. $\square$

It is a strength of the class of HERDU criteria that the two parameters δ and η have a consistent, common interpretation in terms of *intergenerational* inequality aversion. By increasing each of δ and η, inequality aversion is enhanced. For increasing streams, a more inequality averse society discounts the future more, with the social discount rate having a clear ethical significance.

This is in contrast with the class of homothetic DU criteria where the two parameters δ and η represent different ethical notions. The time utility discount rate δ measures the intensity of intergenerational (procedural) *inequity*. A fairer society should choose a lower δ. On the other hand, the elasticity of marginal utility η is often interpreted as a measure of *intra-temporal* inequality aversion. A more egalitarian society should choose a higher η. As a consequence, it is not clear what the social discount rate of an 'equity-minded' society should be: it should discount the future less to avoid intergenerational inequity, but discount the future more because it is more averse to intra-period inequalities.

The result in Proposition 9 has important policy implications, in particular for the question of climate change. If one believes that future generations will be better off in spite of climate change,[5] then a more inequality averse ERDU social observer will agree with the recommendation of Nordhaus [34] to have a gradual emissions-control policy with an increasing carbon price rather than with that of Stern [43], who calls for strong immediate action to mitigate climate change. Indeed, Nordhaus proposes to use $\delta = 0.015$ and $\eta = 2$, whereas Stern argues in favor of $\delta = 0.001$ and $\eta = 1$. However, the policy recommendation will be totally different if one believes that climate change might strongly affect the economy, so that declining consumption would occur for some generations in the future. This perspective may not be unrealistic for some poor developing countries particularly exposed to climate change. In that case, an ERDU social observer using $\eta = 1$ and $\delta > 0$ will recommend discounting future consumption at a negative rate. This rate is lower than the one promoted by Stern for decreasing consumption streams, thus leading to even stronger action.

6. Optimal rank-discounted utilitarian policies

In this section, we establish that ERDU SWOs can be applied to two benchmark models — the Ramsey and Dasgupta–Heal–Solow (DHS) growth models — and show that the ERDU optimal streams in these models are the same as the ones promoted by the SDU SWOs recently studied by Asheim and Mitra [5].

In these models, the ERDU optimal streams maximize DU welfare over all non-decreasing streams. By the justification of sustainability proposed by Asheim, Buchholz and Tungodden [4], Finite Anonymity combined with the Strong Pareto principle rules out all streams that are not non-decreasing when applied to 'productive' technologies. Moreover, ERDU welfare coincides with DU welfare on the set of non-decreasing streams, thereby providing intuition for choosing streams that maximize DU welfare over all non-decreasing streams. However, Asheim, Buchholz and Tungodden's [4] argument is not directly applicable here since (i) ERDU SWOs do not satisfy the Strong Pareto principle for streams that cannot be reordered into non-decreasing streams, and (ii) the DHS growth model is 'productive' only if resource extraction is positive.

For this section, assume that the SWR $\succsim$ on the set of bounded consumption streams is an ERDU SWO represented by W, as defined by Definition 2, where u is assumed to strictly concave and continuously differentiable (on $\mathbb{R}_{++}$) with $\lim_{x\to 0} u'(x) = +\infty$. These additional properties

[5] The assumption is verified in the central scenario of most climate-economy integrated assessment models, such as the RICE model of Nordhaus [34] and the PAGE model used in the Stern Review [43].

on u do not follow from the axiomatic basis for ERDU SWOs, but is imposed on the SWO for the purpose of the analysis of this section. Write $\succsim_{\beta,u}$ for the ERDU SWO determined by β and u, where the properties of u are as described in this paragraph.

Both the Ramsey and DHS models allow for streams that are not bounded above, a complication that must be addressed. For an unbounded stream $\mathbf{x}$, $\ell(\mathbf{x})$ need not exist. If $\ell(\mathbf{x})$ does not exist, then $\mathbf{x}$ can be permuted into a non-decreasing stream, implying that Proposition 1 can be reformulated as follows on any set $\mathbf{X} \subseteq \mathbb{R}_+^{\mathbb{N}}$ where $\mathbf{X}$ admits elements that are not bounded above.

Proposition 1′.

(a) *If $\ell(\mathbf{x})$ does not exist for an allocation $\mathbf{x} \in \mathbf{X}$, then $\mathbf{x}$ belongs to $\overline{\mathbf{X}}$.*
(b) *If $\ell(\mathbf{x})$ exists for an allocation $\mathbf{x} \in \mathbf{X}$ and $|L(\mathbf{x})| < +\infty$, then $\mathbf{x}$ belongs to $\overline{\mathbf{X}}$ if and only if $x_t \leqslant \ell(\mathbf{x})$ for all $t \in \mathbb{N}$.*
(c) *If $\ell(\mathbf{x})$ exists for an allocation $\mathbf{x} \in \mathbf{X}$ and $|L(\mathbf{x})| = +\infty$, then $\mathbf{x}$ belongs to $\overline{\mathbf{X}}$ if and only if $x_t < \ell(\mathbf{x})$ for all $t \in \mathbb{N}$.*

Let W be defined by $W(\mathbf{x}) = \overline{W}(\mathbf{x})$ (cf. Definition 1) if $\ell(\mathbf{x})$ does not exist, while W is defined by Definition 2 if $\ell(\mathbf{x})$ exists. Then $W(\mathbf{x}) = \overline{W}(\bar{\mathbf{x}})$ (cf. Eq. (4)) still holds, where $\bar{\mathbf{x}} = \mathbf{x}$ whenever $\ell(\mathbf{x})$ does not exist.

As shown by Lemmata 1 and 2 in Asheim and Mitra [5], in our applications, the unilateral Laplace transform ($\sum_{t\in\mathbb{N}} \beta^{t-1} x_t$) is finite for any $0 < \beta < 1$ and any feasible $\mathbf{x}$ with these technologies. We will see below that this implies that $W(\mathbf{x})$ is finite for all feasible streams. Hence, let $\mathbf{X} = \{\mathbf{x} \in \mathbb{R}_+^{\mathbb{N}} \mid W(\mathbf{x}) < +\infty\}$ for this section, and let the SWR $\succsim$ on $\mathbf{X}$ be an ERDU SWO represented by W.

The two subsequent subsections introduce sets of feasible streams. A stream $\mathbf{x}$ is *optimal* if $\mathbf{x}$ is feasible and $W(\mathbf{x}) \geqslant W(\mathbf{x}')$ for all feasible streams $\mathbf{x}'$.

6.1. The Ramsey growth model

Assume that the technology is given by a strictly increasing, concave, and continuously differentiable production function $f : \mathbb{R}_+ \to \mathbb{R}_+$, satisfying $f(0) = 0$ and $\lim_{k\to\infty} f'(k) = 0$. A consumption stream $\mathbf{x} = (x_1, x_2, \ldots)$ is feasible given an initial capital stock $k_1 > 0$ if there exists a stream $\{k_2, k_3, \ldots\}$ such that

$$x_t + k_{t+1} \leqslant f(k_t) + k_t, \quad x_t \geqslant 0,\ k_t \geqslant 0, \tag{9}$$

for all $t \in \mathbb{N}$. Such a technology is referred to as a *Ramsey technology*.

Lemma 5. *If $\mathbf{x} \in \mathbf{X}$ is feasible with a Ramsey technology, then $\bar{\mathbf{x}}$ is also feasible. If $\mathbf{x} \in \overline{\mathbf{X}}$ is feasible with a Ramsey technology, then $\mathbf{x}_{[\,]}$ is also feasible.*

Proof. These results follow as storage is costless with a Ramsey technology; cf. Asheim [1, Lemma 3]. □

It follows from Lemma 1 of Asheim and Mitra [5] and Lemma 5 above, combined with the concavity of u, that $W(\mathbf{x})$ is finite for any feasible stream $\mathbf{x}$ with a Ramsey technology.

Following Asheim and Mitra [5] (but changing notation slightly), define the gross output function as $g(k) = f(k) + k$, and denote by $x(y)$ the unique solution to the equation $y = g(y -$

$x(y))$ such that $0 \leqslant x(y) \leqslant y$. The function $x(y)$ is well defined, continuous and differentiable (see [5]). Write $y_\infty(\beta) \equiv \min\{y \geqslant 0 \mid \beta g'(y - x(y)) \leqslant 1\}$. The function y_∞ is strictly increasing for all β for which there exists $k \geqslant 0$ such that $\beta g'(k) = 1$ [5].

Proposition 10. *Consider an ERDU SWO $\succsim_{\beta,u}$ where u is assumed to be strictly concave and continuously differentiable (on $\mathbb{R}_{++}$) with $\lim_{x\to 0} u'(x) = +\infty$, a Ramsey technology, and an initial capital stock $k_1 > 1$. Then there exists a unique optimal consumption stream, denoted by $\mathbf{x}^*$, which is characterized as follows*:

(a) *If $y_1 = g(k_1) \geqslant y_\infty(\beta)$, then $\mathbf{x}^*$ is a stationary stream with $x_t^* = x(y_1)$ for all $t \geqslant 1$.*
(b) *If $y_1 = g(k_1) < y_\infty(\beta)$, then $\mathbf{x}^*$ is an increasing stream, converging to $x(y_\infty(\beta))$ and maximizing $(1-\beta)\sum_{t\in\mathbb{N}} \beta^{t-1} u(x_t)$ over all feasible streams.*

Proof. *Step* 1: *If $\mathbf{x} \in \mathbf{X}$ is optimal, then $\bar{\mathbf{x}} \in \overline{\mathbf{X}}$ is also optimal.* If $\mathbf{x}$ is optimal, then $\mathbf{x}$ is also feasible. By Lemma 5, $\bar{\mathbf{x}}$ is also feasible. By Eq. (4), $W(\bar{\mathbf{x}}) = \overline{W}(\bar{\mathbf{x}}) = W(\mathbf{x}) \geqslant W(\mathbf{x}')$ for all feasible streams $\mathbf{x}'$. Hence, $\bar{\mathbf{x}} \in \overline{\mathbf{X}}$ is also optimal.

Step 2: *If $\mathbf{x} \in \overline{\mathbf{X}}$ is optimal, then $\mathbf{x}_{[\ ]} \in \mathbf{X}^+$ is also optimal.* If $\mathbf{x}$ is optimal, then $\mathbf{x}$ is also feasible. By Lemma 5, $\mathbf{x}_{[\ ]}$ is also feasible. Since $\mathbf{x}_{[\ ]}$ is a permutation of $\mathbf{x}$, $W(\mathbf{x}_{[\ ]}) = W(\mathbf{x}) \geqslant W(\mathbf{x}')$ for all feasible streams $\mathbf{x}'$. Hence, $\mathbf{x}_{[\ ]} \in \mathbf{X}^+$ is also optimal.

Step 3: *If $\mathbf{x} \in \mathbf{X}^+$ is optimal, then $\mathbf{x}$ is the efficient stream $\mathbf{x}^*$ characterized by* (a) *and* (b). The optimality of a non-decreasing $\mathbf{x}$ implies that $\mathbf{x}$ maximizes $\overline{W}(\mathbf{x}')$ over all non-decreasing streams $\mathbf{x}'$. By Proposition 6 of Asheim [1], the efficient stream $\mathbf{x}^*$ characterized by (a) and (b) is the unique stream maximizing $\overline{W}(\mathbf{x}')$ over all non-decreasing streams $\mathbf{x}'$.

Step 4: *If $\mathbf{x} \in \overline{\mathbf{X}} \backslash \mathbf{X}^+$, then $\mathbf{x}$ is not optimal.* Suppose $\mathbf{x} \in \overline{\mathbf{X}} \backslash \mathbf{X}^+$ is optimal. By Step 2, $\mathbf{x}_{[\ ]} \in \mathbf{X}^+$ is optimal. However, by Step 3, if $\mathbf{x}_{[\ ]} \in \mathbf{X}^+$ is optimal, then $\mathbf{x}_{[\ ]}$ coincides with the *efficient* stream $\mathbf{x}^*$ characterized by (a) and (b) However, it is not feasible to permute the efficient $\mathbf{x}^* \in \mathbf{X}^+$ into $\mathbf{x} \in \overline{\mathbf{X}} \backslash \mathbf{X}^+$, as this contradicts the fact that acceleration of consumption along an efficient stream with positive capital stocks is costly with a Ramsey technology [1, Lemma 3]. Hence, $\mathbf{x}$ is not optimal.

Step 5: *If $\mathbf{x} \in \mathbf{X} \backslash \overline{\mathbf{X}}$, then $\mathbf{x}$ is not optimal.* Suppose $\mathbf{x} \in \mathbf{X} \backslash \overline{\mathbf{X}}$ is optimal. By Step 1, $\bar{\mathbf{x}} \in \overline{\mathbf{X}}$ is optimal and, by the property of costless augmentation of initial consumption (cf. [1, Lemma 3]), *inefficient*. However, by Steps 3 and 4, if $\bar{\mathbf{x}} \in \overline{\mathbf{X}}$ is optimal, then $\bar{\mathbf{x}}$ coincides with the *efficient* stream $\mathbf{x}^*$ characterized by (a) and (b). This contradicts that $\mathbf{x}$ is optimal.

Step 6: *The efficient stream $\mathbf{x}^*$ characterized by* (a) *and* (b) *is optimal.* By Proposition 6 of Asheim [1], $W(\mathbf{x}^*) = \overline{W}(\mathbf{x}^*) \geqslant \overline{W}(\mathbf{x}) = W(\mathbf{x})$ if $\mathbf{x}$ is non-decreasing. If $\mathbf{x} \in \overline{\mathbf{X}} \backslash \mathbf{X}^+$, then by Lemma 5, the permutation of $\mathbf{x}$ into the non-decreasing stream $\mathbf{x}_{[\ ]}$ is feasible, and, furthermore, $W(\mathbf{x}^*) = \overline{W}(\mathbf{x}^*) \geqslant \overline{W}(\mathbf{x}_{[\ ]}) = W(\mathbf{x}_{[\ ]}) = W(\mathbf{x})$. Hence, $W(\mathbf{x}^*) \geqslant W(\mathbf{x})$ if $\mathbf{x} \in \overline{\mathbf{X}}$. If $\mathbf{x} \in \mathbf{X} \backslash \overline{\mathbf{X}}$, then $W(\mathbf{x}^*) \geqslant W(\bar{\mathbf{x}}) = \overline{W}(\bar{\mathbf{x}}) = W(\mathbf{x})$ by Eq. (4) since $\bar{\mathbf{x}} \in \overline{\mathbf{X}}$. □

Proposition 10 shows that ERDU preferences can be operationalized in the basic Ramsey model. We are able to characterize a unique optimal solution, which we call the *sustainable discounted utilitarian solution* because it is the same as in Asheim [1] and Asheim and Mitra [5].

Compared to SDU preferences, ERDU preferences emphasize more clearly the influence of inequality aversion on optimal policy. Indeed, we know that a necessary condition for an ERDU SWF $W_{\beta,u}$ to be more inequality averse than another ERDU SWF $W_{\hat{\beta},\hat{u}}$ is that $\beta \leqslant \hat{\beta}$. From Proposition 10, it follows that:

- A more inequality averse ERDU society $\succsim_{\beta,u}$ converges to a lower steady state consumption than a less inequality averse society $\succsim_{\hat{\beta},\hat{u}}$ whenever k_1 satisfies $g(k_1) < y_\infty(\hat{\beta})$ and leads to the same steady state consumption otherwise.
- A more inequality averse ERDU society $\succsim_{\beta,u}$ prevents growth for a larger set of initial conditions than a less inequality averse society $\succsim_{\hat{\beta},\hat{u}}$ (for k_1 satisfying $g(k_1) \geqslant y_\infty(\beta)$, as opposed to k_1 satisfying $g(k_1) \geqslant y_\infty(\hat{\beta})$).

Regarding the second point, recall that maximin always prevents growth. Maximin is the special case of ERDU preferences in which $\beta \to 0$, an extreme aversion to inequality. However, growth is also prevented for low values of β.

Inequality aversion therefore modifies both the long-run perspectives of the society and the prospects of an egalitarian (stationary) distribution. Only the parameter β determines the long-term impact of inequality aversion. The other dimension of inequality aversion, the concavity of the function u, has only an impact on the speed of the convergence to the steady state when $g(k_1) < y_\infty(\beta)$.

6.2. The Dasgupta–Heal–Solow growth model

The Dasgupta–Heal–Solow model [17,42] is the standard model of growth with an exhaustible natural resource. Production depends on man-made physical capital k_t^m, on the extraction d_t of a natural exhaustible resource k_t^n and on the labor supply ℓ_t. The natural resource is depleted by the resource use, so that $k_{t+1}^n = k_t^n - d_t$. The production function $\hat{f}(k_t^m, d_t, l_t)$ is concave, non-decreasing, homogeneous of degree one, and twice continuously differentiable. It satisfies $(\hat{f}_{k^m}, \hat{f}_d, \hat{f}_\ell) \gg 0$ for all $(k^m, d, \ell) \gg 0$ and $\hat{f}(k^m, 0, \ell) = \hat{f}(0, d, \ell) = 0$ (both physical capital and the natural resource are essential in production). Moreover, given $(\tilde{k}^m, \tilde{d}) \gg 0$, there exists a scalar $\tilde{\chi}$ such that $(d\hat{f}_d(k^m, d, 1))/(\hat{f}_\ell(k^m, d, 1)) \geqslant \tilde{\chi}$ for (k^m, d) satisfying $k^m \geqslant \tilde{k}^m$ and $0 \leqslant d \leqslant \tilde{d}$ (the ratio of the share of the resource to the share of labor is bounded away from zero when labor is fixed at a unit level).

Assume that the labor force is constant and normalized to 1. Write $f(k^m, d) := \hat{f}(k^m, d, 1)$. Also assume that f is strictly concave and $f_{k^m,d}(k^m, d) \gg 0$ for all $(k^m, d) \gg 0$. A consumption stream $\mathbf{x} = (x_1, x_2, \ldots)$ is feasible given initial stocks $(k_1^m, k_1^n) \gg 0$ if there exists a stream $\{(k_2^m, k_2^n), (k_3^m, k_3^n), \ldots\}$ such that

$$x_t + k_{t+1}^m \leqslant f\left(k_t^m, k_t^n - k_{t+1}^n\right) + k_t^m, \quad x_t \geqslant 0,\ k_t^m \geqslant 0,\ k_t^n \geqslant 0, \tag{10}$$

for all $t \in \mathbb{N}$. Hence, production $f(k_t^m, k_t^n - k_{t+1}^n)$ is split between consumption x_t and capital accumulation $k_{t+1}^m - k_t^m$ at each time t.

The assumptions made so far do not ensure that it is feasible to maintain a constant and positive consumption level forever. Therefore, assume in addition that there exists from any $(k_1^m, k_1^n) \gg 0$, a constant stream with positive consumption. Cass and Mitra [13] give a necessary and sufficient condition on f for this assumption to hold. Under this additional assumption there exists an *efficient* constant consumption stream from any $(k_1^m, k_1^n) \gg 0$ (see [19, Proposition 5]). A technology satisfying the above assumptions is referred to as a *Dasgupta–Heal–Solow* (DHS) *technology*.

When establishing the implications of ERDU SWOs in the DHS growth model, it is a complication that production is increasing in capital only if resource extraction is positive, but a constant function of capital if resource extraction is zero. However, the analysis of the Ramsey model above can still be adapted to the DHS growth model.

Lemma 6. *If* $\mathbf{x} \in \mathbf{X}$ *is feasible with a DHS technology, then* $\bar{\mathbf{x}}$ *is also feasible. If* $\mathbf{x} \in \overline{\mathbf{X}}$ *is feasible with a DHS technology, then* $\mathbf{x}_{[\,]}$ *is also feasible.*

Proof. These results follow as storage is costless with a DHS technology; cf. Asheim [1, Lemma 4]. □

It follows from Lemma 2 of Asheim and Mitra [5] and Lemma 6 above, combined with the concavity of u, that $\overline{W}(\mathbf{x})$ is finite for any feasible stream $\mathbf{x}$ with a DHS technology.

Denote by $x(k_1^m, k_1^n)$ the positive and constant level of consumption that can be sustained forever along an efficient constant consumption stream from $(k_1^m, k_1^n) \gg 0$. It is possible to attach a sequence of shadow prices

$$\big(p_1(k_1^m, k_1^n), p_2(k_1^m, k_1^n), \ldots, p_t(k_1^m, k_1^n), \ldots\big)$$

to the corresponding stationary consumption stream (for a characterization of the prices, see [5, Lemma 3]). Write

$$\beta_\infty(k_1^m, k_1^n) = \frac{\sum_{t=2}^{+\infty} p_t(k_1^m, k_1^n)}{\sum_{t=1}^{+\infty} p_t(k_1^m, k_1^n)}$$

for the long-run discount factor at time 1 supporting this stationary stream.

Proposition 11. *Consider an ERDU SWF* $\succsim_{\beta,u}$ *where* u *is assumed to be strictly concave and continuously differentiable (on* $\mathbb{R}_{++}$*) with* $\lim_{x\to 0} u'(x) = +\infty$*, a Dasgupta–Heal–Solow technology, and initial stocks and resource* $(k_1^m, k_1^n) \gg 0$*. Then there exists a unique optimal consumption stream, denoted by* $\mathbf{x}^*$*, which is characterized as follows*:

(a) *If* $\beta_\infty(k_1^m, k_1^n) \geqslant \beta$*, then* $\mathbf{x}^*$ *is a stationary stream with* $x_t^* = x(k_1^m, k_1^n)$ *for all* $t \geqslant 1$.
(b) *If* $\beta_\infty(k_1^m, k_1^n) < \beta$*, then* $\mathbf{x}^*$ *is a non-decreasing stream maximizing* $(1-\beta)\sum_{t\in\mathbb{N}} \beta^{t-1} u(x_t)$ *over all feasible and non-decreasing streams. The stream exhibits the following pattern*:
   - *For* $t < \tau$, $x_t^* < x_{t+1}^*$,
   - *for all* $t \geqslant \tau$, $x_t^* = x((k_\tau^m)^*, (k_\tau^n)^*)$,

   where $\tau := \min\{t \in \mathbb{N} \mid \beta_\infty((k_t^m)^*, (k_t^n)^*) \geqslant \beta\}$.

The proof of Proposition 11 closely follows the proof of Proposition 10, differing only by substituting Lemma 6 for Lemma 5, Asheim and Mitra [5, Lemma 6] for Asheim [1, Proposition 6], and Asheim [1, Lemma 4] for Asheim [1, Lemma 3]. Hence, it is not repeated here.

Proposition 11 shows that the consequences of a higher level of inequality aversion exhibited in the Ramsey growth model still hold in the Dasgupta–Heal–Solow model. Indeed:

- A more inequality averse ERDU society $\succsim_{\beta,u}$ will prevent growth for a larger set of initial conditions than a less inequality averse society $\succsim_{\hat{\beta},\hat{u}}$ (for (k_1^m, k_1^n) satisfying $\beta_\infty(k_1^m, k_1^n) \geqslant \beta$, as opposed to (k_1^m, k_1^n) satisfying $\beta_\infty(k_1^m, k_1^n) \geqslant \hat{\beta}$).

In particular, in the maximin case, growth is always prevented. Again, the maximin case represents an extreme form of inequality aversion, and less extreme degrees of inequality aversion may allow for growth in an initial phase.

7. Conclusion

The ERDU approach to intertemporal welfare has several appealing features. First, it offers a continuous and numerically representable criterion that reconciles intergenerational procedural equity and efficiency on the set of allocations that can be rearranged into non-decreasing streams. Second, compared to procedurally equitable, but incomplete, criteria like undiscounted utilitarianism and lexicographic maximin, it allows for more flexibility in the specification of inequality aversion. This is of particular importance in conflicts between the present generation and an infinite number of future generations (cf. [2]). Last, it provides a consistent and intuitive interpretation of the ethical parameters determining the social discount rate. With the ERDU interpretation, we have obtained the provocative statement that inequality aversion increases the social discount rate along increasing consumption streams.

This statement is at odds with the traditional ethical approach to social discounting. It comes from the fact that ERDU criteria satisfy procedural equity (the reason why people endorsing the traditional ethical approach have called for lower discount rates) while allowing for inequality-aversion-based discounting. We believe that ERDU may spark off new debates on social discounting within the ethical approach to social discounting.

The ERDU criterion can be operationalized. In particular, in benchmark growth models, the ERDU optimal policies coincide with those promoted by the sustainable discounted utilitarian criterion that has been recently studied by Asheim and Mitra [5]. While its recommendations may not be new, the ERDU criterion offers an interesting new perspective that respects procedural equity and displays concerns for intergenerational redistribution. It sheds some new lights on what the present generation owes to future generations. On the one hand, we must guarantee that they will not be worse off than we are. On the other hand, intergenerational inequalities in favor of future generations should not be too large as this would be unfair to the present generation. This conception of intergenerational equity, more in line with the intuitive notion of distributional equity, may seem appealing to many.

Appendix A

Proof of Proposition 2. Assume that $\succsim$ satisfies Order, Continuity, Monotonicity, Restricted Dominance, Restricted Separable Present, Restricted Separable Future, Restricted Stationarity and Strong Anonymity. Order, Continuity, and Monotonicity imply that there exists a monotonic SWF $\overline{W}$ representing $\succsim$ on $\overline{\mathbf{X}}$. By Strong Anonymity, for all $\mathbf{x} \in \overline{\mathbf{X}}$, $\overline{W}(\mathbf{x}) = \overline{W}(\mathbf{x}_{[\,]})$. We can therefore restrict attention to the set $\mathbf{X}^+$.

Now, for each $T \in \mathbb{N}$, we introduce the following subset of $\mathbf{X}^+$:

$$\left\{\mathbf{x} \in \mathbf{X}^+\colon\ x_t = x_{T+1},\ \forall t \geqslant T+1\right\}.$$

These are the non-decreasing intergenerational allocations with a constant tail from period $T+1$ onward. Denote the restriction of $\succsim$ to this set by $\succsim_T$, which is a continuous monotonic weak order on the following rank-ordered set:

$$\mathbf{X}_\mathbf{T}^+ = \left\{(x_1, \dots, x_{T+1}) \in \mathbb{R}^{T+1}\colon\ x_1 \leqslant \dots \leqslant x_{T+1}\right\}.$$

Let $\mathcal{T} = \{1, \dots, T+1\}$ denote the set of indices of the coordinates of $\mathbf{X}_\mathbf{T}^+$.

We then proceed by showing that $\succsim_T$ satisfies a separability property. A subset of coordinates $\mathcal{S} \subset \mathcal{T}$ is said to be *separable* for $\succsim_T$ if for all $(x_1, \dots, x_{T+1})$, $(y_1, \dots, y_{T+1})$, $(x'_1, \dots, x'_{T+1})$,

$(y'_1, \ldots, y'_{T+1})$ in $\mathbf{X}_\mathbf{T}^+$, if $x_s = x'_s$ and $y_s = y'_s$ for all $s \in \mathcal{S}$ and $x_t = y_t$ and $x'_t = y'_t$ for all $t \in \mathcal{T} \setminus \mathcal{S}$, then:

$$(x_1, \ldots, x_{T+1}) \succsim_T (y_1, \ldots, y_{T+1}) \quad \Leftrightarrow \quad \left(x'_1, \ldots, x'_{T+1}\right) \succsim_T \left(y'_1, \ldots, y'_{T+1}\right).$$

A subset of coordinates $\mathcal{S} \subset \mathcal{T}$ is said to be *essential* if there exist $(v_s)_{s\in\mathcal{S}}$ and $(w_s)_{s\in\mathcal{S}}$ in $\mathbb{R}^{|\mathcal{S}|}$ and $(z_t)_{t\in\mathcal{T}\setminus\mathcal{S}}$ in $\mathbb{R}^{|\mathcal{T}\setminus\mathcal{S}|}$ such that, if $(x_1, \ldots, x_{T+1})$ and $(y_1, \ldots, y_{T+1})$ are defined by $x_s = v_s$ and $y_s = w_s$ for all $s \in \mathcal{S}$ and $x_t = y_t = z_t$ for all $t \in \mathcal{T} \setminus \mathcal{S}$, then $(x_1, \ldots, x_{T+1}) \in \mathbf{X}_\mathbf{T}^+$, $(y_1, \ldots, y_{T+1}) \in \mathbf{X}_\mathbf{T}^+$ and $(x_1, \ldots, x_{T+1}) \succ_T (y_1, \ldots, y_{T+1})$. The set $\mathcal{T}$ is completely separable for the relation $\succsim_T$ if every subset $\mathcal{S} \subset \mathcal{T}$ is separable and essential.

To show that the set $\mathcal{T} = \{1, \ldots, T+1\}$ is completely separable for the ordering $\succsim_T$, we use Theorem 1 in Gorman [26]. Let $(\mathcal{I}, \mathcal{K}, \mathcal{L}, \mathcal{M})$ be a partition of the set of indices $\mathcal{T}$ such that each subset is not the empty set. The theorem states that: If $\mathcal{I} \cup \mathcal{K}$ and $\mathcal{L} \cup \mathcal{K}$ are separable and essential for $\succsim_T$, then $\mathcal{I}$, $\mathcal{K}$, $\mathcal{L}$, $\mathcal{I} \cup \mathcal{L}$ and $\mathcal{I} \cup \mathcal{K} \cup \mathcal{L}$ are separable and essential for $\succsim_T$.

For any $t \leqslant T$, consider the following elements of $\mathbf{X}_\mathbf{T}^+$:

$$(z_1, \ldots, z_{t-1}, x_t, x_{t+1}, z_{t+2}, \ldots, z_{T+1}), \quad (z_1, \ldots, z_{t-1}, y_t, y_{t+1}, z_{t+2}, \ldots, z_{T+1}),$$
$$(w_1, \ldots, w_{t-1}, x_t, x_{t+1}, w_{t+2}, \ldots, w_{T+1}), \quad (w_1, \ldots, w_{t-1}, y_t, y_{t+1}, w_{t+2}, \ldots, w_{T+1}).$$

By repeated application of Restricted Separable Future and Restricted Stationarity, we have that:

$$\begin{aligned}&(z_1, \ldots, z_{t-1}, x_t, x_{t+1}, z_{t+2}, \ldots, z_{T+1}) \succsim_T (z_1, \ldots, z_{t-1}, y_t, y_{t+1}, z_{t+2}, \ldots, z_{T+1}) \\ &\quad \Leftrightarrow \quad (x_t, x_{t+1}, z_{t+2}, \ldots, z_{T+1}) \succsim_{T-t+1} (y_t, y_{t+1}, z_{t+2}, \ldots, z_{T+1}).\end{aligned}$$

By Restricted Separable Present,

$$\begin{aligned}&(x_t, x_{t+1}, z_{t+2}, \ldots, z_{T+1}) \succsim_{T-t+1} (y_t, y_{t+1}, z_{t+2}, \ldots, z_{T+1}) \\ &\quad \Leftrightarrow \quad (x_t, x_{t+1}, w_{t+2}, \ldots, w_{T+1}) \succsim_{T-t+1} (y_t, y_{t+1}, w_{t+2}, \ldots, w_{T+1}).\end{aligned}$$

And once again by repeated application of Restricted Separable Future and Restricted Stationarity,

$$\begin{aligned}&(x_t, x_{t+1}, w_{t+2}, \ldots, w_{T+1}) \succsim_{T-t+1} (y_t, y_{t+1}, w_{t+2}, \ldots, w_{T+1}) \\ &\quad \Leftrightarrow \quad (w_1, \ldots, w_{t-1}, x_t, x_{t+1}, w_{t+2}, \ldots, w_{T+1}) \\ &\qquad\qquad \succsim_T (w_1, \ldots, w_{t-1}, y_t, y_{t+1}, w_{t+2}, \ldots, w_{T+1}).\end{aligned}$$

As a consequence,

$$\begin{aligned}&(z_1, \ldots, z_{t-1}, x_t, x_{t+1}, z_{t+2}, \ldots, z_{T+1}) \succsim_T (z_1, \ldots, z_{t-1}, y_t, y_{t+1}, z_{t+2}, \ldots, z_{T+1}) \\ &\quad \Leftrightarrow \quad (w_1, \ldots, w_{t-1}, x_t, x_{t+1}, w_{t+2}, \ldots, w_{T+1}) \\ &\qquad\qquad \succsim_T (w_1, \ldots, w_{t-1}, y_t, y_{t+1}, w_{t+2}, \ldots, w_{T+1}),\end{aligned}$$

establishing that the set $\{t, t+1\}$ is separable for $\succsim_T$ for all $t \leqslant T$.

By Restricted Dominance and repeated application of Restricted Separable Future and Restricted Stationarity as above, the set $\{t, t+1\}$ is also essential.

Setting $\mathcal{I} = \{t\}$, $\mathcal{K} = \{t+1\}$ and $\mathcal{L} = \{t+2\}$, by Gorman's theorem the sets $\{t\}$, $\{t+1\}$, $\{t+2\}$, $\{t, t+2\}$ and $\{t, t+1, t+2\}$ are separable and essential. Repeating this reasoning, all sets $\{t, t'\} \subset \mathcal{T}$ are separable and essential. By taking unions of such sets and Gorman's theorem, we can obtain any subset $\mathcal{S} \subset \mathcal{T}$.[6] Hence, $\mathcal{T} = \{1, \dots, T+1\}$ is completely separable for the relation $\succsim_T$.

Because $\mathbf{X}_\mathbf{T}^+$ is a rank-ordered set and $\mathcal{T} = \{1, \dots, T+1\}$ is completely separable for the relation $\succsim_T$, we know by Theorem 3.2 and Corollary 3.6 of Wakker [47] that there exists a cardinal additive representation of $\succsim_T$:

$$\overline{W}_T(\mathbf{x}) = \sum_{t=1}^{T} u_{t,T}(x_t) + V_T(x_{T+1}), \quad \forall \mathbf{x} \in \mathbf{X}_\mathbf{T}^+. \tag{11}$$

The functions $u_{t,T}$ and V_T are all continuous and non-decreasing. In addition, by Monotonicity, Restricted Dominance, Restricted Separable Future and Restricted Stationarity, the functions $u_{1,T}$ and V_T must be increasing. By cardinality, we may set $u_{t,T}(0) = 0$ for all $t \leqslant T$ and $V_T(0) = 0$ (the normalization condition).

Now, representation (11) exists for $\succsim_T$ for all $T \in \mathbb{N}$. Furthermore, $\succsim_T$ and $\succsim_{T+1}$ represent the same ordering on $\mathbf{X}_\mathbf{T}^+$. By standard uniqueness results for additive functions on rank-ordered sets [47, Theorem 3.2], we can let (after the appropriate normalization) $u_{t,T} \equiv u_{t,T+1}$ and $V_T \equiv u_{T,T+1} + V_{T+1}$. We can henceforth drop the subscript T in the functions $u_{t,T}$.

By Restricted Separable Future and Restricted Stationarity, we also know that $\overline{W}_T(\mathbf{x}) = \sum_{t=1}^{T} u_t(x_t) + V_T(x_{T+1})$ and $\overline{W}_T(\mathbf{x}) = \sum_{t=2}^{T+1} u_t(x_{t-1}) + V_{T+1}(x_{T+1})$ represent the same preferences for all $\mathbf{x} \in \mathbf{X}_\mathbf{T}^+$. By the cardinality of the additive representation and the normalization condition, there must exist a $\beta > 0$ such that $u_{t+1}(x) = \beta u_t(x)$ and $V_{T+1}(x) = \beta V_T(x)$ for all $x \in \mathbb{R}_+$. Note that β does not depend on t. Let $\bar{u} \equiv u_1$ and $V \equiv V_1$, we have the following representation of $\succsim_T$:

$$\overline{W}_T(\mathbf{x}) = \sum_{t=1}^{T} \beta^{t-1} \bar{u}(x_t) + \beta^T V(x_{T+1}), \quad \forall \mathbf{x} \in \mathbf{X}_\mathbf{T}^+,$$

with $\bar{u}$ and V being two increasing functions.

Now note that we must also have $V(x) = \bar{u}(x) + \beta V(x)$, so that $V(x) = \bar{u}(x)/(1-\beta)$. This implies that $\beta < 1$ by Monotonicity and Restricted Dominance. This also implies that $V(x) = \sum_{t=1}^{+\infty} \beta^{t-1} \bar{u}(x)$. Hence, we obtain the following representation of $\succsim$ on $\bigcup_{T \in \mathbb{N}} \mathbf{X}_\mathbf{T}^+$ by letting $u(x) \equiv \bar{u}(x)/(1-\beta)$:

$$\overline{W}(\mathbf{x}) = (1-\beta) \sum_{t=1}^{+\infty} \beta^{t-1} u(x_t).$$

Now it remains to prove that the representation extends to the whole set $\mathbf{X}^+$. For any $\mathbf{x} \in \mathbf{X}^+$, we define the sequence $\mathbf{x}^1, \mathbf{x}^2, \dots, \mathbf{x}^k, \dots$ of allocations in $\mathbf{X}^+$ as follows: for any $k \in \mathbb{N}$, $x_t^k = x_t$ for all $t \leqslant k$ and $x_t^k = x_k$ for all $t > k$. Each allocation in the sequence belongs to $\bigcup_{T \in \mathbb{N}} \mathbf{X}_\mathbf{T}^+$, and $\lim_{k \to \infty} \sup_{t \in \mathbb{N}} |x_t^k - x_t| = 0$ because we consider bounded streams. By Continuity, we obtain that $\overline{W}(\mathbf{x}) = (1-\beta) \sum_{t=1}^{+\infty} \beta^{t-1} u(x_t)$ is an SWF representing $\succsim$ on $\mathbf{X}^+$. □

Proof of Lemma 3. We prove the result for the case where $T = 0$. The extension to the case where $T > 0$ (provided that $|L(\mathbf{x})| \geqslant T$) is straightforward: pull these dates out and redo the arguments below on the remainder of the stream.

[6] Singleton sets having been obtained in the preceding steps.

Hence, by Eq. (2) we seek to establish that, for all $\mathbf{x} \in \mathbf{X}$,

$$\widetilde{W}(\mathbf{x}) \leqslant u\big(\ell(\mathbf{x})\big).$$

Let $\mathbf{y}$ be defined by, $\forall t \in \mathbb{N}$, $y_t = \max\{x_t, \ell(\mathbf{x})\} \geqslant x_t$. By Monotonicity, $\widetilde{W}(\mathbf{x}) \leqslant \widetilde{W}(\mathbf{y})$. Hence, it is sufficient to show that $\widetilde{W}(\mathbf{y}) \leqslant u(\ell(\mathbf{x}))$. Write $m := \sup_t y_t$. Note that $m \in \mathbb{R}_+$ exists since $\mathbf{y}$ is bounded.

Case 1: $m = \ell(\mathbf{x})$. Then $y_t = \ell(\mathbf{x})$ for all t and $\widetilde{W}(\mathbf{y}) = u(\ell(\mathbf{x}))$.

Case 2: $m > \ell(\mathbf{x})$. W.l.o.g. normalize the consumption scale s.t. $m = 1$ and $\ell(\mathbf{x}) = 0$. Because $\ell(\mathbf{x}) = 0$, we can construct

$$\mathbf{z} = \left(1, \dots, 1, \frac{1}{2}, 1, \dots, 1, \frac{1}{3}, 1, \dots, 1, \frac{1}{4}, 1, \dots, 1, \frac{1}{5}, 1, \dots, 1, \frac{1}{6}, 1, \dots\right)$$

by starting with 1 and continuing with 1 until there is some τ such that $\frac{1}{2} \geqslant y_\tau$; going back to 1 and continuing with 1 until there is some τ' such that $\frac{1}{3} \geqslant y_{\tau'}$; and so on. By the definition of $m = 1$, $\forall t \in \mathbb{N}$, $z_t \geqslant y_t$. By Monotonicity, $\widetilde{W}(\mathbf{x}) \leqslant \widetilde{W}(\mathbf{y}) \leqslant \widetilde{W}(\mathbf{z})$. Hence, it is sufficient to show that $\widetilde{W}(\mathbf{z}) \leqslant u(0)$.

By Strong Anonymity, $\mathbf{z}$ is indifferent to each member of the following sequence of streams (where in $\mathbf{z}^i$, the 1s appear at even dates t satisfying $t \geqslant 2i$):

$$\mathbf{z}^1 = \left(\frac{1}{2}, 1, \frac{1}{3}, 1, \frac{1}{4}, 1, \frac{1}{5}, 1, \frac{1}{6}, 1, \dots\right),$$
$$\mathbf{z}^2 = \left(\frac{1}{3}, \frac{1}{2}, \frac{1}{4}, 1, \frac{1}{5}, 1, \frac{1}{6}, 1, \frac{1}{7}, 1, \dots\right),$$
$$\mathbf{z}^3 = \left(\frac{1}{4}, \frac{1}{3}, \frac{1}{5}, \frac{1}{2}, \frac{1}{6}, 1, \frac{1}{7}, 1, \frac{1}{8}, 1, \dots\right),$$
$$\mathbf{z}^4 = \left(\frac{1}{5}, \frac{1}{4}, \frac{1}{6}, \frac{1}{3}, \frac{1}{7}, \frac{1}{2}, \frac{1}{8}, 1, \frac{1}{9}, 1, \dots\right),$$
$$\vdots$$

Write, $\forall i, j \in \mathbb{N}$, $m^{ij} := \max_{t \in \{1,\dots,j\}} z_t^i$. If $j \geqslant 2i$, then $m^{ij} = 1$. If there exists $k \in \mathbb{Z}_+$ such that $j < 2(i-k)$, then $m^{ij} \leqslant 1/(2+k)$. Define, $\forall i, j \in \mathbb{N}$, $\mathbf{w}^{ij}$ by, $\forall t \in \mathbb{N}$, $w_t^{ij} = m^{ij}$ if $t \leqslant j$ and $w_t^{ij} = 1$ otherwise.

Since, $\forall i, j \in \mathbb{N}$, $\mathbf{w}^{ij} \in \mathbf{X}^+$ and, $\forall t \in \mathbb{N}$, $w_t^{ij} \geqslant z_t^i$, it follows that, for all $i, j \in \mathbb{N}$,

$$\widetilde{W}(\mathbf{x}) \leqslant \widetilde{W}(\mathbf{y}) \leqslant \widetilde{W}(\mathbf{z}) = \widetilde{W}\big(\mathbf{z}^i\big) \leqslant \widetilde{W}\big(\mathbf{w}^{ij}\big) = (1-\beta)\sum_{t\in\mathbb{N}} \beta^{t-1} u\big(w_t^{ij}\big)$$

because $\succsim$ satisfies Monotonicity and Strong Anonymity and is represented on $\overline{\mathbf{X}} \supseteq \mathbf{X}^+$ by $\overline{W}$.

Suppose $\widetilde{W}(\mathbf{z}) = u(0) + \epsilon$, where $\epsilon > 0$. Since $0 < \beta < 1$, one can choose $j \in \mathbb{N}$ such that $\beta^j(u(1) - u(0)) < \frac{1}{2}\epsilon$. Since u is continuous and, for fixed $j \in \mathbb{N}$, $m^{ij} \to 0$ as $i \to \infty$, one can choose $i \in \mathbb{N}$ such that $(1-\beta^j)u(m^{ij}) < (1-\beta^j)u(0) + \frac{1}{2}\epsilon$. Then

$$\widetilde{W}\big(\mathbf{w}^{ij}\big) = \big(1-\beta^j\big)u\big(m^{ij}\big) + \beta^j u(1) < u(0) + \frac{1}{2}\epsilon + \frac{1}{2}\epsilon = u(0) + \epsilon.$$

This contradicts that $\widetilde{W}(\mathbf{z}) \leqslant \widetilde{W}(\mathbf{w}^{ij})$ for all $i, j \in \mathbb{N}$. Hence, $\widetilde{W}(\mathbf{x}) \leqslant \widetilde{W}(\mathbf{y}) \leqslant \widetilde{W}(\mathbf{z}) \leqslant u(0) = u(\ell(\mathbf{x}))$. $\square$

Proof of Lemma 4. Define $\mathbf{y}$ by, $\forall t \in \mathbb{N}$, $y_t = \ell(\mathbf{x})$. Construct a sequence of streams, $\mathbf{y}^j$, $j \in \mathbb{Z}_+$, inductively as follows: $\mathbf{y}^0 = \mathbf{y}$ and, $\forall j \in \mathbb{N}$,

$$\mathbf{y}^j = \begin{cases} \mathbf{y}^{j-1} & \text{if } x_j \geqslant \ell(\mathbf{x}), \\ (y_1^{j-1}, \ldots, y_{j-1}^{j-1}, x_j, {}_{j+1}\mathbf{y}^{j-1}) & \text{if } x_j < \ell(\mathbf{x}). \end{cases}$$

Note that, $\forall j \in \mathbb{N}$, $\ell(\mathbf{y}^j) = \ell(\mathbf{x})$. By Lemma 3, $\forall j \in \mathbb{N}$,

$$\widetilde{W}(\mathbf{x}) \leqslant \overline{W}\big(\mathbf{y}^j\big). \tag{12}$$

Case 1: $|L(\mathbf{x})| < \infty$. In this case, there exists $j \in \mathbb{N}$ s.t., $\mathbf{y}^j = \bar{\mathbf{x}}$. As $\mathbf{y}^j \in \overline{\mathbf{X}}$, $\overline{W}(\mathbf{y}^j) = \widetilde{W}(\mathbf{y}^j) \leqslant \widetilde{W}(\mathbf{x})$ by Monotonicity. Hence, by Eq. (4),

$$\widetilde{W}(\mathbf{x}) = \overline{W}\big(\mathbf{y}^j\big) = \overline{W}(\bar{\mathbf{x}}) = W(\mathbf{x}).$$

Case 2: $|L(\mathbf{x})| = \infty$. Define, $\forall j \in \mathbb{N}$, $\ell^j := \min_{t>j} x_t$. Note that $\ell^j \to \ell(\mathbf{x})$ as $j \to \infty$. For each $j \in \mathbb{N}$, define $\mathbf{z}^j$ by, $\forall t \in \mathbb{N}$, $z_t^j = \min\{x_t, \ell^j\}$. By Monotonicity, $\forall j \in \mathbb{N}$, it holds that $\overline{W}(\mathbf{z}^j) = \widetilde{W}(\mathbf{z}^j) \leqslant \widetilde{W}(\mathbf{x})$ because $\mathbf{z}^j \in \overline{\mathbf{X}}$ and $\ell^j < \ell(\mathbf{x})$. Hence, by (12) it follows that, $\forall j \in \mathbb{N}$,

$$\overline{W}\big(\mathbf{z}^j\big) \leqslant \widetilde{W}(\mathbf{x}) \leqslant \overline{W}\big(\mathbf{y}^j\big). \tag{13}$$

For any $\epsilon > 0$, we can choose $j \in \mathbb{N}$ s.t. $\overline{W}(\mathbf{y}^j) - \overline{W}(\mathbf{z}^j) \leqslant u(\ell(\mathbf{x})) - u(\ell^j) < \epsilon$ since u is continuous and $\ell^j \to \ell(\mathbf{x})$ as $j \to \infty$. Combined with (13) and Eq. (4), this implies $\widetilde{W}(\mathbf{x}) = \lim_{j\to\infty} \overline{W}(\mathbf{y}^j) = \overline{W}(\bar{\mathbf{x}}) = W(\mathbf{x})$. □

Proof of Proposition 6. *If the ERDU SWO $\succsim_{\beta,u}$ satisfies Pigou–Dalton Transfer, then $\beta \times \mathcal{C}_u \leqslant 1$.* Assume that the ERDU SWO $\succsim_{\beta,u}$ satisfies Pigou–Dalton Transfer. Consider $\mathbf{x} \in \mathbf{X}$ such that $x_t = 0$ for all $t \leqslant \tau$, $x_\tau \leqslant x_{\tau+1}$, and $x_t > x_\tau + x_{\tau+1}$ for $t > \tau + 1$. Now consider $\mathbf{y} \in \mathbf{X}$ such that $y_\tau + \varepsilon = x_\tau \leqslant x_{\tau+1} = y_{\tau+1} - \varepsilon$ and $y_t = x_t$ for all $t \neq \tau, \tau+1$, with $x_\tau \geqslant \varepsilon > 0$. Since the ERDU SWO $\succsim_{\beta,u}$ satisfies Pigou–Dalton Transfer, it follows from the representation (3) of ERDU SWOs that:

$$\beta^{\tau-1}u(x_\tau) + \beta^\tau u(x_{\tau+1}) \geqslant \beta^{\tau-1}u(y_\tau) + \beta^\tau u(y_{\tau+1}) = \beta^{\tau-1}u(x_\tau - \varepsilon) + \beta^\tau u(x_{\tau+1} + \varepsilon).$$

This inequality can be rewritten as

$$1 \geqslant \beta \frac{u(x_{\tau+1} + \varepsilon) - u(x_{\tau+1})}{u(x_\tau) - u(x_\tau - \varepsilon)}. \tag{14}$$

The construction of allocations $\mathbf{x}$ and $\mathbf{y}$ yielding this inequality is possible for any two integers $\tau < \tau'$ and for any real numbers $0 < \varepsilon \leqslant x_\tau \leqslant x_{\tau'}$. Hence,

$$1 \geqslant \sup_{0<\varepsilon\leqslant x\leqslant x'} \beta \times \frac{u(x' + \varepsilon) - u(x')}{u(x) - u(x - \varepsilon)} = \beta \times \mathcal{C}_u.$$

If $\beta \times \mathcal{C}_u \leqslant 1$ holds for the ERDU SWO $\succsim_{\beta,u}$, then $\succsim_{\beta,u}$ satisfies Pigou–Dalton Transfer. Assume that $1 \geqslant \beta \times \mathcal{C}_u$ holds for the ERDU SWO $\succsim_{\beta,u}$. Consider $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ such that $\varepsilon \leqslant y_\tau + \varepsilon = x_\tau \leqslant x_{\tau'} = y_{\tau'} - \varepsilon$ and $y_t = x_t$ for all $t \neq \tau, \tau'$, where $x_\tau \geqslant \varepsilon > 0$ and $\tau, \tau' \in \mathbb{N}$. We want to show that $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$.

If $\ell(\mathbf{x}) \leqslant x_{\tau'}$, then $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$ by a dominance argument for those components that matter. In the case where $x_{\tau'} < \ell(\mathbf{x}) < y_{\tau'}$, the same argument as below applies, replacing $y_{\tau'}$ by $\ell(\mathbf{x})$. Therefore, assume that $y_{\tau'} \leqslant \ell(\mathbf{x})$.

Using representation (3), it follows that $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$ if and only if:

$$W(\mathbf{x}) - W(\mathbf{y}) = \sum_{\bar{r}_\tau(\mathbf{y}) \leqslant r \leqslant r_\tau(\mathbf{x})} \beta^{r-1}\big(u(x_{[r]}) - u(y_{[r]})\big) - \sum_{\bar{r}_{\tau'}(\mathbf{x}) \leqslant r \leqslant r_{\tau'}(\mathbf{y})} \beta^{r-1}\big(u(y_{[r]}) - u(x_{[r]})\big) \geqslant 0.$$

For $\bar{r}_\tau(\mathbf{y}) \leqslant r \leqslant r_\tau(\mathbf{x})$, we have $u(x_{[r]}) - u(y_{[r]}) \geqslant 0$, and for $\bar{r}_{\tau'}(\mathbf{x}) \leqslant r \leqslant r_{\tau'}(\mathbf{y})$, we have $u(y_{[r]}) - u(x_{[r]}) \geqslant 0$. Hence:

$$\sum_{\bar{r}_\tau(\mathbf{y}) \leqslant r \leqslant r_\tau(\mathbf{x})} \beta^{r-1}\big(u(x_{[r]}) - u(y_{[r]})\big) - \sum_{\bar{r}_{\tau'}(\mathbf{x}) \leqslant r \leqslant r_{\tau'}(\mathbf{y})} \beta^{r-1}\big(u(y_{[r]}) - u(x_{[r]})\big)$$
$$\geqslant \beta^{r_\tau(\mathbf{x})-1} \times \sum_{\bar{r}_\tau(\mathbf{y}) \leqslant r \leqslant r_\tau(\mathbf{x})} \big(u(x_{[r]}) - u(y_{[r]})\big) - \beta^{\bar{r}_{\tau'}(\mathbf{x})-1} \times \sum_{\bar{r}_{\tau'}(\mathbf{x}) \leqslant r \leqslant r_{\tau'}(\mathbf{y})} \big(u(y_{[r]}) - u(x_{[r]})\big).$$

By the definition of the Pigou–Dalton Transfer, $\sum_{\bar{r}_\tau(\mathbf{y}) \leqslant r \leqslant r_\tau(\mathbf{x})}(u(x_{[r]}) - u(y_{[r]})) = u(x_\tau) - u(y_\tau)$ and $\sum_{\bar{r}_{\tau'}(\mathbf{x}) \leqslant r \leqslant r_{\tau'}(\mathbf{y})}(u(y_{[r]}) - u(x_{[r]})) = u(y_{\tau'}) - u(x_{\tau'})$. Therefore:

$$\begin{aligned} W(\mathbf{x}) - W(\mathbf{y}) &\geqslant \beta^{r_\tau(\mathbf{x})-1}\big(u(x_\tau) - u(y_\tau)\big) - \beta^{\bar{r}_{\tau'}(\mathbf{x})-1}\big(u(y_{\tau'}) - u(x_{\tau'})\big) \\ &= \beta^{r_\tau(\mathbf{x})-1}\big(u(x_\tau) - u(y_\tau)\big)\left(1 - \beta^{\bar{r}_{\tau'}(\mathbf{x}) - r_\tau(\mathbf{x})} \frac{u(y_{\tau'}) - u(x_{\tau'})}{u(x_\tau) - u(y_\tau)}\right) \\ &\geqslant \beta^{r_\tau(\mathbf{x})-1}\big(u(x_\tau) - u(y_\tau)\big)\left(1 - \beta \times \frac{u(x_{\tau'} + \varepsilon) - u(x_{\tau'})}{u(x_\tau) - u(x_\tau - \varepsilon)}\right). \end{aligned}$$

Since $1 \geqslant \beta \times \mathcal{C}_u$, it follows that $W(\mathbf{x}) - W(\mathbf{y}) \geqslant 0$ and thus $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$. □

The next lemma is needed for the proof of Proposition 7. Write

$$\hat{X} = \big\{(x_1, x_2, x_3, x_4) \in \mathbb{R}^4_+ \colon 0 \leqslant x_1 < x_2 \leqslant x_3 < x_4, \text{ where } x_4 = \hat{u}^{-1}\big(\hat{u}(x_3) + \hat{\beta}^{\tau-\tau'}\big(\hat{u}(x_2) - \hat{u}(x_1)\big)\big) \text{ for some } \tau, \tau' \in \mathbb{N} \text{ with } \tau < \tau'\big\}$$

and

$$\widetilde{\mathcal{C}}_{u,\hat{u}} = \sup_{(x_1,x_2,x_3,x_4) \in \hat{X}} \frac{[u(x_4) - u(x_3)]/[\hat{u}(x_4) - \hat{u}(x_3)]}{[u(x_2) - u(x_1)]/[\hat{u}(x_2) - \hat{u}(x_1)]}.$$

Lemma 7. $\mathcal{C}_{u,\hat{u}} = \widetilde{\mathcal{C}}_{u,\hat{u}}$.

Proof. Let $y_1 = \hat{u}(x_1)$, $y_2 = \hat{u}(x_2)$, $y_3 = \hat{u}(x_3)$ and $y_4 = \hat{u}(x_4)$. Then

$$\mathcal{C}_{u,\hat{u}} = \sup_{0 \leqslant y_1 < y_2 \leqslant y_3 < y_4} \frac{[u \circ \hat{u}^{-1}(y_4) - u \circ \hat{u}^{-1}(y_3)]/[y_4 - y_3]}{[u \circ \hat{u}^{-1}(y_2) - u \circ \hat{u}^{-1}(y_1)]/[y_2 - y_1]} = G_{u \circ \hat{u}^{-1}},$$

where $G_{u \circ \hat{u}^{-1}}$ is Chateauneuf, Cohen and Meilijson's [14] 'greediness' index for the function $u \circ \hat{u}^{-1}$. Also let

$$\hat{X}_\lambda = \left\{(y_1, y_2, y_3, y_4) \in \mathbb{R}^4_+ \colon 0 \leqslant y_1 < y_2 \leqslant y_3 < y_4, \text{ where } \frac{y_4 - y_3}{y_2 - y_1} = \lambda\right\}$$

and

$$G_{u\circ\hat{u}^{-1}}(\lambda) = \sup_{(y_1,y_2,y_3,y_4)\in\hat{X}_\lambda} \frac{[u\circ\hat{u}^{-1}(y_4) - u\circ\hat{u}^{-1}(y_3)]/[y_4 - y_3]}{[u\circ\hat{u}^{-1}(y_2) - u\circ\hat{u}^{-1}(y_1)]/[y_2 - y_1]}.$$

Then $\widetilde{\mathcal{C}}_{u,\hat{u}} = \sup_{\lambda=\hat{\beta}^{\tau-\tau'},\tau<\tau'} G_{u\circ\hat{u}^{-1}}(\lambda)$. By Chateauneuf, Cohen and Meilijson [14, Lemma 1], $G_{u\circ\hat{u}^{-1}} = G_{u\circ\hat{u}^{-1}}(\lambda)$ for any $\lambda > 0$. Hence, $\mathcal{C}_{u,\hat{u}} = \widetilde{\mathcal{C}}_{u,\hat{u}}$. □

Proof of Proposition 7. *If $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$, then $\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}$.* Assume that $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$. Consider $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ such that, for some $\tau, \tau' \in \mathbb{N}$ with $\tau < \tau'$:

- $x_t = y_t = 0$ for all $t < \tau$;
- $0 \leqslant y_\tau < x_\tau \leqslant y_t = x_t \leqslant x_{\tau'}$ for all $\tau < t < \tau'$;
- $y_{\tau'} = \hat{u}^{-1}(\hat{u}(x_{\tau'}) + \hat{\beta}^{\tau-\tau'}(\hat{u}(x_\tau) - \hat{u}(y_\tau)))$, so that $y_{\tau'} \geqslant x_{\tau'}$;
- $y_t = x_t > y_\tau$ for $t > \tau'$.

(a) By construction, $\hat{\beta}^{\tau'}(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'})) = \hat{\beta}^{\tau}(\hat{u}(x_\tau) - \hat{u}(y_\tau))$, so $\mathbf{x} \sim_{\hat{\beta},\hat{u}} \mathbf{y}$.

(b) Because $\mathbf{y} \succ_I \mathbf{x}$ and $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$, $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$, it follows that $\beta^{\tau'}(u(y_{\tau'}) - u(x_{\tau'})) \leqslant \beta^{\tau}(u(x_\tau) - u(y_\tau))$. Facts (a) and (b) imply

$$\frac{\beta^\tau/\hat{\beta}^\tau}{\beta^{\tau'}/\hat{\beta}^{\tau'}} \geqslant \frac{(u(y_{\tau'}) - u(x_{\tau'}))/(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'}))}{(u(x_\tau) - u(y_\tau))/(\hat{u}(x_\tau) - \hat{u}(y_\tau))}.$$

The construction of $\mathbf{x}$ and $\mathbf{y}$ yielding this inequality is possible for any two integers $\tau < \tau'$ and for any real numbers $0 \leqslant y_\tau < x_\tau \leqslant x_{\tau'} < y_{\tau'}$ such that $y_{\tau'} = \hat{u}^{-1}(\hat{u}(x_{\tau'}) + \hat{\beta}^{\tau-\tau'}(\hat{u}(x_\tau) - \hat{u}(y_\tau)))$. Then

$$\mathcal{D}_{\beta,\hat{\beta}} = \inf_{t<t'} \frac{\beta^\tau/\hat{\beta}^\tau}{\beta^{\tau'}/\hat{\beta}^{\tau'}} \geqslant \sup_{(x_1,x_2,x_3,x_4)\in\hat{X}} \frac{[u(x_4) - u(x_3)]/[\hat{u}(x_4) - \hat{u}(x_3)]}{[u(x_2) - u(x_1)]/[\hat{u}(x_2) - \hat{u}(x_1)]} = \widetilde{\mathcal{C}}_{u,\hat{u}}.$$

By Lemma 7 it follows that $\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}$.

If $\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}$, then $\succsim_{\beta,u}$ is at least as inequality averse as $\succsim_{\hat{\beta},\hat{u}}$. Assume that $\mathbf{y} \succ_I \mathbf{x}$ and $\mathbf{x} \sim_{\hat{\beta},\hat{u}} \mathbf{y}$.[7] We want to show that $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$ if $\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}$.

If $\ell(\mathbf{x}) \leqslant x_{\tau'}$, then $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$ and $\mathbf{x} \succsim_{\hat{\beta},\hat{u}} \mathbf{y}$. If $x_{\tau'} < \ell(\mathbf{x}) < y_{\tau'}$, then the argument below applies, replacing $y_{\tau'}$ by $\ell(\mathbf{x})$. Therefore, assume that $y_{\tau'} \leqslant \ell(\mathbf{x})$.

By using $\hat{\beta}^{r_\tau(\mathbf{x})}(\hat{u}(x_\tau) - \hat{u}(y_\tau)) = \hat{\beta}^{\bar{r}_{\tau'}(\mathbf{x})}(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'}))$ (which follows from Eq. (3) and $\mathbf{x} \sim_{\hat{\beta},\hat{u}} \mathbf{y}$), we obtain

$$\begin{aligned}
&\beta^{r_\tau(\mathbf{x})}\big(u(x_\tau) - u(y_\tau)\big) - \beta^{\bar{r}_{\tau'}(\mathbf{x})}\big(u(y_{\tau'}) - u(x_{\tau'})\big) \\
&\quad = \hat{\beta}^{\bar{r}_{\tau'}(\mathbf{x})}\big(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'})\big)\left(\frac{\beta^{r_\tau(\mathbf{x})}(u(x_\tau) - u(y_\tau))}{\hat{\beta}^{r_\tau(\mathbf{x})}(\hat{u}(x_\tau) - \hat{u}(y_\tau))} - \frac{\beta^{\bar{r}_{\tau'}(\mathbf{x})}(u(y_{\tau'}) - u(x_{\tau'}))}{\hat{\beta}^{\bar{r}_{\tau'}(\mathbf{x})}(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'}))}\right)
\end{aligned}$$

[7] If $\mathbf{x} \succ_{\hat{\beta},\hat{u}} \mathbf{y}$, then by monotonicity there exists $x_\tau - y_\tau > \varepsilon > 0$ such that $\mathbf{x}'$ defined by $x'_\tau = x_\tau - \varepsilon$ (where τ is as in Definition 4) and $x'_t = x_t$ for all $t \neq \tau$ satisfies $\mathbf{x}' \sim_{\hat{\beta},\hat{u}} \mathbf{y}$. By monotonicity, $\mathbf{x} \succ_{\beta,u} \mathbf{x}'$. By transitivity, $\mathbf{x} \succ_{\beta,u} \mathbf{y}$. It is also the case that $\mathbf{y} \succ_I \mathbf{x}'$. Hence, if $(\mathbf{y} \succ_I \mathbf{x} \,\&\, \mathbf{x} \sim_{\hat{\beta},\hat{u}} \mathbf{y}) \Rightarrow (\mathbf{x} \succsim_{\beta,u} \mathbf{y})$, then $(\mathbf{y} \succ_I \mathbf{x} \,\&\, \mathbf{x} \succ_{\hat{\beta},\hat{u}} \mathbf{y}) \Rightarrow (\mathbf{x} \succ_{\beta,u} \mathbf{y})$.

$$= \frac{\beta^{\bar{r}_{\tau'}(\mathbf{x})}(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'}))}{\frac{\hat{u}(x_\tau)-\hat{u}(y_\tau)}{u(x_\tau)-u(y_\tau)}}\left(\left(\frac{\beta}{\hat{\beta}}\right)^{r_\tau(\mathbf{x})-\bar{r}_{\tau'}(\mathbf{x})} - \frac{\frac{u(y_{\tau'})-u(x_{\tau'})}{\hat{u}(y_{\tau'})-\hat{u}(x_{\tau'})}}{\frac{u(x_\tau)-u(y_\tau)}{\hat{u}(x_\tau)-\hat{u}(y_\tau)}}\right)$$

$$\geqslant \frac{\beta^{\bar{r}_{\tau'}(\mathbf{x})}(\hat{u}(y_{\tau'}) - \hat{u}(x_{\tau'}))}{\frac{\hat{u}(x_\tau)-\hat{u}(y_\tau)}{u(x_\tau)-u(y_\tau)}}(\mathcal{D}_{\beta,\hat{\beta}} - \mathcal{C}_{u,\hat{u}}).$$

It now follows from Eq. (3) that $\mathbf{x} \succsim_{\beta,u} \mathbf{y}$ whenever $\mathcal{D}_{\beta,\hat{\beta}} \geqslant \mathcal{C}_{u,\hat{u}}$. □

References

[1] G.B. Asheim, Unjust intergenerational allocations, J. Econ. Theory 54 (1991) 350–371.
[2] G.B. Asheim, Intergenerational equity, Annu. Rev. Econ. 2 (2010) 197–222.
[3] G.B. Asheim, W. Buchholz, The malleability of undiscounted utilitarianism as a criterion of intergenerational justice, Economica 70 (2003) 405–422.
[4] G.B. Asheim, W. Buchholz, B. Tungodden, Justifying sustainability, J. Environ. Econ. Manage. 41 (2001) 252–268.
[5] G.B. Asheim, T. Mitra, Sustainability and discounted utilitarianism in models of economic growth, Math. Soc. Sci. 59 (2010) 148–169.
[6] G.B. Asheim, T. Mitra, B. Tungodden, Sustainable recursive social welfare functions, Econ. Theory (2010), doi:10.1007/s00199-010-0573-7, forthcoming.
[7] A.B. Atkinson, On the measurement of inequality, J. Econ. Theory 2 (1970) 244–263.
[8] K. Basu, T. Mitra, Aggregating infinite utility streams with intergenerational equity: The impossibility of being Paretian, Econometrica 71 (2003) 1557–1563.
[9] C.R. Birchenhall, P. Grout, On equal plans with infinite horizon, J. Econ. Theory 21 (1979) 249–264.
[10] H. Bleichrodt, K.I.M. Rhode, P.P. Wakker, Koopmans' constant discounting for intertemporal choice: A simplification and a generalization, J. Math. Psych. 52 (2008) 341–347.
[11] W. Bossert, An axiomatization of the single-series Ginis, J. Econ. Theory 50 (1990) 82–92.
[12] W. Bossert, Y. Sprumont, K. Suzumura, Ordering infinite utility streams, J. Econ. Theory 135 (2007) 579–589.
[13] D. Cass, T. Mitra, Indefinitely sustained consumption despite exhaustible natural resources, Econ. Theory 1 (1991) 119–146.
[14] A. Chateauneuf, M. Cohen, I. Meilijson, More pessimism than greediness: A characterization of monotone risk aversion in the rank-dependent expected utility model, Econ. Theory 25 (2005) 649–667.
[15] G. Chichilnisky, An axiomatic approach to sustainable development, Soc. Choice Welfare 13 (1996) 231–257.
[16] P.S. Dasgupta, Discounting climate change, J. Risk Uncertainty 37 (2008) 141–169.
[17] P.S. Dasgupta, G. Heal, The optimal depletion of exhaustible resources, in: Symposium on the Economics of Exhaustible Resources, Rev. Econ. Stud. 41 (1974) 3–28.
[18] P.S. Dasgupta, G. Heal, Economic Theory and Exhaustible Resources, Cambridge Univ. Press, Cambridge, UK, 1979.
[19] S. Dasgupta, T. Mitra, Intergenerational equity and efficient allocation of exhaustible resources, Int. Econ. Rev. 24 (1983) 133–153.
[20] P. Diamond, The evaluation of infinite utility streams, Econometrica 33 (1965) 170–177.
[21] D. Donaldson, J.A. Weymark, A single-parameter generalization of the Gini indices of inequality, J. Econ. Theory 22 (1980) 67–86.
[22] D. Donaldson, J.A. Weymark, Ethically flexible Gini indices for income distributions in the continuum, J. Econ. Theory 29 (1983) 353–358.
[23] U. Ebert, Measurement of inequality: An attempt at unification and generalization, Soc. Choice Welfare 5 (1988) 147–169.
[24] M. Fleurbaey, P. Michel, Intertemporal equity and the extension of the Ramsey criterion, J. Math. Econ. 39 (2003) 777–802.
[25] I. Gilboa, Expected utility with purely subjective non-additive probabilities, J. Math. Econ. 16 (1987) 65–88.
[26] W.G. Gorman, The structure of utility functions, Rev. Econ. Stud. 35 (1968) 367–390.
[27] S. Grant, J. Quiggin, Increasing uncertainty: A definition, Math. Soc. Sci. 49 (2005) 117–141.
[28] C. Hara, T. Shinotsuka, K. Suzumura, Y. Xu, Continuity and egalitarianism in the evaluation of infinite streams, Soc. Choice Welfare 31 (2008) 179–191.
[29] R.F. Harrod, Towards a Dynamic Economy: Some Recent Developments of Economic Theory and Their Application to Policy, Macmillan, London, 1948.

[30] T.C. Koopmans, Stationary ordinal utility and impatience, Econometrica 28 (1960) 287–309.
[31] L. Lauwers, Ordering infinite utility streams comes at the cost of a non-Ramsey set, J. Math. Econ. 46 (2010) 32–37.
[32] J.A. Mirrlees, Optimum growth when technology is changing, Rev. Econ. Stud. 34 (1967) 95–124.
[33] W.D. Nordhaus, The Stern review on the economics of climate change, J. Econ. Lit. 45 (2007) 687–702.
[34] W.D. Nordhaus, A Question of Balance: Weighing the Options on Global Warming Policies, Yale Univ. Press, New Haven, CT, 2008.
[35] A.C. Pigou, The Economics of Welfare, Macmillan, London, 1920.
[36] J. Quiggin, A theory of anticipated utility, J. Econ. Behav. Organ. 3 (1982) 323–343.
[37] F.P. Ramsey, A mathematical theory of saving, Econ. J. 38 (1928) 543–559.
[38] J. Rawls, A Theory of Justice, revised edition, The Belknap Press of the Harvard Univ. Press, Cambridge, MA, 1999.
[39] Y. Rébillé, Patience in some non-additive models, J. Math. Econ. 43 (2007) 749–763.
[40] T. Sakai, An impartial representation in infinite time horizon, mimeo, Keio University, 2011.
[41] D. Schmeidler, Subjective probability and expected utility without additivity, Econometrica 57 (1989) 571–587.
[42] R.M. Solow, Intergenerational equity and exhaustible resources, in: Symposium on the Economics of Exhaustible Resources, Rev. Econ. Stud. 41 (1974) 29–45.
[43] N.H. Stern, The Stern Review of the Economics of Climate Change, Cambridge Univ. Press, Cambridge, UK, 2006.
[44] L.-G. Svensson, Equity among generations, Econometrica 48 (1980) 1251–1256.
[45] L. Van Liedekerke, L. Lauwers, Sacrificing the patrol: Utilitarianism, future generations and infinity, Econ. Philos. 13 (1997) 159–174.
[46] P.P. Wakker, Additive representations on rank-ordered sets I. The algebraic approach, J. Math. Psych. 35 (1991) 501–531.
[47] P.P. Wakker, Additive representations on rank-ordered sets II. The topological approach, J. Math. Econ. 22 (1993) 1–26.
[48] M.L. Weitzman, A review of the Stern review of the economics of climate change, J. Econ. Lit. 45 (2007) 703–724.
[49] J.A. Weymark, Generalized Gini inequality indices, Math. Soc. Sci. 1 (1981) 409–430.
[50] M.E. Yaari, The dual theory of choice under risk, Econometrica 55 (1987) 95–115.
[51] W.R. Zame, Can intergenerational equity be operationalized?, Theoretical Econ. 2 (2007) 187–202.