

Asheim, Geir B.; Zuber, Stéphane

Working Paper

Probability Adjusted Rank-Discounted Utilitarianism

CESifo Working Paper, No. 4728

Provided in Cooperation with:

Ifo Institute – Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Asheim, Geir B.; Zuber, Stéphane (2014) : Probability Adjusted Rank-Discounted Utilitarianism, CESifo Working Paper, No. 4728, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:

<https://hdl.handle.net/10419/96833>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



Working Papers

www.cesifo.org/wp

Probability Adjusted Rank-Discounted Utilitarianism

Geir B. Asheim
Stéphane Zuber

CESIFO WORKING PAPER NO. 4728
CATEGORY 2: PUBLIC CHOICE
MARCH 2014

An electronic version of the paper may be downloaded

- *from the SSRN website:* www.SSRN.com
- *from the RePEc website:* www.RePEc.org
- *from the CESifo website:* www.CESifo-group.org/wp

Probability Adjusted Rank-Discounted Utilitarianism

Abstract

We propose and axiomatize probability adjusted rank-discounted critical-level generalized utilitarianism (PARDCLU). We thus generalize rank-discounted utilitarianism (RDU) (proposed by Zuber and Asheim, 2012) to variable population and risky situations and thereby take important steps towards preparing RDU for practical use, e.g. for evaluation of climate policies and other policy issues with long-run consequences. We illustrate how PARDCLU yields rank-dependent expected utilitarianism - but with additional structure - in a special case, and show how PARDCLU can handle a situation with positive probability of human extinction.

JEL-Code: D630, D710, D810, H430, Q560.

Keywords: social evaluation, population ethics, decision-making under risk, critical-level utilitarianism, social discounting.

Geir B. Asheim
Department of Economics
University of Oslo
Norway
g.b.asheim@econ.uio.no

Stéphane Zuber
Paris School of Economics & CNRS
Centre d'Économie de la Sorbonne
Paris / France
stephane.zuber@univ-paris1.fr

March 6, 2014

We thank Reyer Gerlagh for helpful discussions and seminar and conference participants at Lund, Hitotsubashi, Oslo and Toulouse Universities for valuable suggestions. 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. Asheim's research has also been supported by l'Institut d'études avancées - Paris. Zuber's research has been supported by the Chair on Welfare Economics and Social Justice at the Institute for Global Studies (FMSH - Paris) and the Franco-Swedish Program on Economics and Philosophy (FMSH & Riksbankens Jubileumsfond).

1 Introduction

Economic evaluation of climate change is usually based on discounted utilitarianism, where transformed wellbeing (utility) is discounted by a constant and positive per-period rate and summed over all time periods. As a matter of principle, such utility discounting means that generations are treated unequally (at least if the utility discount rate exceeds the per-period probability of human extinction). Moreover, as a matter of practical policy evaluation, this criterion is virtually insensitive to the long-term effects of climate change, beyond year 2100 when the most serious consequences are expected to occur.

Equal treatment of generations is often associated with *undiscounted utilitarianism*, where utility is summed without discounting. However, when modeling the many potential future people by assuming that there are infinitely many generations, this criterion assigns zero relative weight to the present generation's interests. It leads to the unappealing prescription that the present generation should endure heavy sacrifices even if it contributes to only a tiny gain for all future generations.

The criterion of maximizing the wellbeing of the worst-off generation (*maximin*) also treats generations equally, but assigns zero relative weight to all generations but the worst-off. It leads to the unappealing prescription that the present generation should not do an even negligible sacrifice for the benefit of better off future generations.

This dilemma—that the ethically commendable requirement of equal treatment seems to lead to extreme prescriptions when applied in actual criteria of intergenerational equity—is a main motivation for rank-discounted utilitarianism (RDU), proposed and analyzed by Zuber and Asheim (2012). RDU discounts future utility as long as the future is better than the present, thereby trading-off current sacrifice and future gain. However, if the present generation is better off than all future generations, then priority shifts to the future. In this case, zero relative weight is assigned to present utility. The criterion of RDU can therefore capture the intuition that we should be more willing to assist future generations if they are worse off than

us, but not to save much for their benefit if they turn out to be better off. RDU is compatible with equal treatment of generations as discounting is made according to rank, not according to time.

Applying RDU to a distributional problem like climate change requires that the criterion explicitly takes into account that population size depends endogenously on the chosen policy and that there are uncertain future consequences of present policy. Moreover, the issues of population and risk might be interrelated. In particular, there might be a positive probability of human extinction.

In Asheim and Zuber (2013) we contribute to population ethics by proposing and axiomatizing *rank-discounted critical-level generalized utilitarianism* (RDCLU). Moreover, we establish how RDCLU avoids serious objections raised against other variable population criteria. In particular, it escapes both

- the *Repugnant Conclusion* (Parfit, 1976, 1982, 1984) where, for any population with excellent lives, there is a population with lives barely worth living that is better, provided that the latter includes sufficiently many people
- the *Very Sadistic Conclusion* (Arrhenius, 2000, 2012) where, for any population with terrible lives not worth living, there is a population with good lives that is worse, provided that the latter includes sufficiently many people.

In the present paper we extend RDCLU to risky situations, including the case with positive probability of human extinction, by proposing the *probability adjusted rank-discounted critical-level generalized utilitarian* (PARDCLU) social welfare order (Definition 1). We start out in Section 2 by presenting a framework where each (potential) individual is characterized by a level of lifetime wellbeing and a probability of existence. We illustrate how this set-up can be derived from a formulation where information arrives in each of T time periods, with individuals living for one period only and not being subjected to risk during their lifetime, reflecting an intergenerational perspective.

We then, in Section 3, present an axiomatic foundation for PARDCLU through Theorem 1, which is proven in the appendix. A key axiom, called *Probability adjusted*

Suppes-Sen, combines the strict Pareto principle with anonymity. In conjunction with the Continuity axiom, it implies invariance to permutations of individuals with the same wellbeing and the same probability of existence. It also entails invariance to the replacement of one individual with given wellbeing and probability with two individuals having the same wellbeing and whose probability of existence sum up the probability of original individual.

In the subsequent Section 4 we illustrate the usefulness of PARDCLU by showing its consequences in various special cases. In the special case where the individual probabilities of existence sum up to one, PARDCLU yields rank-dependent expected utilitarianism, but with additional structure. This additional structure derives from the axiom *Existence independence of the worst-off*, which plays the same role as Koopmans' (1960) stationarity postulate. We also show how PARDCLU handles human extinction. In Sections 5–7 we discuss some issues faced by the PARDCLU approach, in particular how it can be practically implemented by dealing with the planning horizon, how it may achieve time consistency, and how it relates to the ex ante vs. ex post debate in social choice theory. In the final Section 8 we provide concluding remarks.

2 Framework

Let \mathbb{N} denote the natural numbers, let \mathbb{R} denote the real numbers, let \mathbb{R}_+ (resp. \mathbb{R}_{++}) denote the non-negative (resp. positive) real numbers, and let \mathbb{Q}_{++} denote the positive rational numbers.

Individuals are described by two numbers: their lifetime wellbeing and their probability of existence. An allocation $\mathbf{x} \in (\mathbb{R} \times (0, 1])^n$ determines the finite population size, $n(\mathbf{x}) = n$, and the distribution of wellbeing and probability,

$$\mathbf{x} = (x_1, \dots, x_{n(\mathbf{x})}) = \left((x_1^w, x_1^p), \dots, (x_{n(\mathbf{x})}^w, x_{n(\mathbf{x})}^p) \right),$$

among the $n(\mathbf{x})$ individuals that make up the population. For each $i \in \{1, \dots, n(\mathbf{x})\}$, x_i^w is the individual's wellbeing and x_i^p is her probability of existence. We denote

by $\nu(\mathbf{x}) = \sum_{i=1}^{n(\mathbf{x})} x_i^p$ the probability adjusted population size of \mathbf{x} and by

$$\mathbf{X} = \cup_{n \in \mathbb{N}} (\mathbb{R} \times (0, 1])^n$$

the set of possible finite *allocations*.

The concept of an allocation, as defined above, can be derived from a formulation where information arrives in each of T time periods. To see this, let S be the finite set of potential signals in each $t = 1, \dots, T$, let $\Omega = S^T$ denote the set of states of the world, and let $p \in \Delta(\Omega)$ be an objective probability distribution over this set.

For any $\omega = (s^1, \dots, s^T) \in \Omega$ and $t = 1, \dots, T$, let $\omega^t = (s^1, \dots, s^t) \in S^t$ be the history of signals up to time t , with $p^t(\omega^t)$ denoting the probability of ω^t derived from p . For each $t = 1, \dots, T$, let the function $n^t : S^t \rightarrow \mathbb{N} \cup \{0\}$ determine the number of individuals living at time t , and let the function $w^t : \{\omega^t \in S^t : n^t(\omega^t) > 0\} \rightarrow \cup_{n \in \mathbb{N}} \mathbb{R}^n$ determine the distribution of wellbeing after histories at time t with positive population size. For each $t = 1, \dots, T$ and $\omega^t \in S^t$ with $n^t(\omega^t) > 0$, individual j 's pair of wellbeing and probability of existence, where $j \in \{1, \dots, n^t(\omega^t)\}$, is given by $(w_j^t(\omega^t), p^t(\omega^t))$. Hence an allocation \mathbf{x} correspond to a probability distribution p , where $\mathbf{x} = (((w_j^t(\omega^t), p^t(\omega^t))_{j \in \{1, \dots, n^t(\omega^t)\}})_{\omega^t \in S^t})_{t \in \{1, \dots, T\}}$. The functions n^t and w^t are permitted to vary so that we can retrieve the whole set \mathbf{X} from this dynamic setting although the set of signals is finite.

This formulation implies that individuals live for one period only, and are not subjected to risk during their lifetime. Rather, there is social risk associated with the lifetime wellbeing of future individuals. Our focus on intergenerational issues motivates this abstraction from lifetime fluctuations and individual risk.

Following the usual convention in population ethics, lifetime wellbeing equal to 0 represents *neutrality*. Hence, lifetime wellbeing is normalized so that above neutrality, a life, as a whole, is worth living; below neutrality, it is not.

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

For each $\mathbf{x} \in \mathbf{X}$, let $\pi : \{1, \dots, n(\mathbf{x})\} \rightarrow \{1, \dots, n(\mathbf{x})\}$ be a bijection that reorders individuals in increasing wellbeing order:

$$x_{\pi(r)}^w \leq x_{\pi(r+1)}^w \quad \text{for all } r \in \{1, \dots, n(\mathbf{x}) - 1\}.$$

Let $\rho_0 = 0$ and define the probability adjusted rank ρ_r inductively as follows:

$$\rho_r = x_{\pi(r)}^p + \rho_{r-1}$$

for $r \in \{1, \dots, n(\mathbf{x})\}$. Define the *rank-ordered allocation* $\mathbf{x}_{[\cdot]} : (0, \nu(\mathbf{x})) \rightarrow \mathbb{R}$ by

$$\mathbf{x}_{[\rho]} = x_{\pi(r)}^w \quad \text{for } \rho_{r-1} < \rho \leq \rho_r \text{ and } 1 < r \leq n(\mathbf{x})$$

and write $\mathbf{x}_{[0]} := \lim_{\rho \downarrow 0} \mathbf{x}_{[\rho]}$. Note that the permutation π need not be unique (if, for instance, $x_i^w = x_{i'}^w$ for some $i \neq i'$), but the resulting rank-ordered allocation $\mathbf{x}_{[\cdot]}$ is unique. Note also that the definitions imply that $\rho_{n(\mathbf{x})} = \nu(\mathbf{x})$.

For every $\nu \in \mathbb{R}$, write $\mathbf{X}_\nu = \{\mathbf{x} \in \mathbf{X} : \nu(\mathbf{x}) = \nu\}$ for the set of finite allocations with probability adjusted population size equal to ν . For $\mathbf{x}, \mathbf{y} \in \mathbf{X}_\nu$, write $\mathbf{x}_{[\cdot]} > \mathbf{y}_{[\cdot]}$ if $\mathbf{x}_{[\rho]} \geq \mathbf{y}_{[\rho]}$ for all $\rho \in (0, \nu]$ and $\mathbf{x}_{[\rho']} > \mathbf{y}_{[\rho']}$ for some $\rho' \in (0, \nu]$; note that, by the definitions of the step functions $\mathbf{x}_{[\cdot]}$ and $\mathbf{y}_{[\cdot]}$, $\mathbf{x}_{[\rho']} > \mathbf{y}_{[\rho']}$ implies that $\mathbf{x}_{[\rho]} > \mathbf{y}_{[\rho]}$ for all ρ in a subset of $(0, \nu]$ that includes a non-empty proper interval.

For $z \in \mathbb{R}$, $p \in (0, 1]$ and $n \in \mathbb{N}$, let $\mathbf{x} \in (\mathbb{R} \times (0, 1])^n$ with $(x_i^w, x_i^p) = (z, p)$ for all $i \in \{1, \dots, n\}$ be denoted by $(z)_\nu$, where $\nu = np$. For $\mathbf{x} \in \mathbf{X}$, $z \in \mathbb{R}$, $p \in (0, 1]$ and $n \in \mathbb{N}$, let $\mathbf{y} \in (\mathbb{R} \times (0, 1])^{n(\mathbf{x})+n}$ such that $(y_i^w, y_i^p) = (x_i^w, x_i^p)$ for all $i \in \{0, \dots, n(\mathbf{x})\}$ and $(y_i^w, y_i^p) = (z, p)$ for all $i \in \{n(\mathbf{x}) + 1, \dots, n(\mathbf{x}) + n\}$ be denoted by $(\mathbf{x}, (z)_{np})$.

3 Axioms and representation result

Probability adjusted rank-discounted critical-level utilitarianism can be characterized by the following seven axioms.

The first three axioms are sufficient to ensure numerical representation of the SWR for any fixed probability adjusted population size. They also entail that individuals are treated anonymously and with sensitivity for their well-being.

Axiom 1 (Order) *The relation \succsim is complete, reflexive and transitive on \mathbf{X} .*

An SWR satisfying Axiom 1 is called a social welfare order (SWO).

Axiom 2 (Continuity) *For all $\nu \in \mathbb{R}_{++}$ and $\mathbf{x} \in \mathbf{X}_\nu$, the sets $\{\mathbf{y} \in \mathbf{X}_\nu : \mathbf{y} \succsim \mathbf{x}\}$ and $\{\mathbf{y} \in \mathbf{X} : \mathbf{x} \succsim \mathbf{y}\}$ are closed for the topology induced by the supnorm applied to rank-ordered allocations.¹*

Axiom 3 (Probability adjusted Suppes-Sen) *For all $\nu \in \mathbb{R}_{++}$ and $\mathbf{x}, \mathbf{y} \in \mathbf{X}_\nu$, if $\mathbf{x}_{[\cdot]} > \mathbf{y}_{[\cdot]}$, then $\mathbf{x} \succ \mathbf{y}$.*

Jointly with axiom 2, axiom 3 implies anonymity wrt. different individuals with the same probability of existence. Hence, permuting the wellbeing levels of individuals with the same probability of existence leads to an equally good allocation.

In line with the Asheim and Zuber's (2013) axiomatization of rank-discounted critical-level utilitarianism we impose independence to adding an individual only if the added individual is best-off (relative to two allocations with the same probability adjusted population size) or worst-off.

Axiom 4 (Existence independence of the best-off) *For all $\nu \in \mathbb{R}_{++}$, $\mathbf{x}, \mathbf{y} \in \mathbf{X}_\nu$, $p \in (0, 1]$, and $z \in \mathbb{R}$ satisfying $z \geq \max\{\mathbf{x}_{[\nu]}, \mathbf{y}_{[\nu]}\}$, $(\mathbf{x}, (z)_p) \succsim (\mathbf{y}, (z)_p)$ if and only if $\mathbf{x} \succsim \mathbf{y}$.*

Axiom 5 (Existence independence of the worst-off) *For all $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, $p \in (0, 1]$, and $z \in \mathbb{R}$ satisfying $z \leq \min\{\mathbf{x}_{[0]}, \mathbf{y}_{[0]}\}$, $(\mathbf{x}, (z)_p) \succsim (\mathbf{y}, (z)_p)$ if and only if $\mathbf{x} \succsim \mathbf{y}$.*

Moreover, we introduce a critical wellbeing level $c \in \mathbb{R}_+$, which if experienced by an added individual without changing the utilities of the existing population, leads

¹This means that we use the metric $d(x, y) = \sup_{r \in [0, \nu]} |x_{[r]} - y_{[r]}|$. In functional spaces, the topology induced by the sup metric is strong, so that the associated notion of continuity is weak (in the sense that if continuity would hold for a weaker topology, it would be satisfied for the stronger topology proposed here). This is an advantage of our definition.

to an alternative which is as good as the original if $\mathbf{x}_{[\nu(\mathbf{x})]} \leq c$. Since $c \geq 0$, c is at least as large as the neutral wellbeing level.

Axiom 6 (Existence of a critical level) *There exist $c \in \mathbb{R}_+$ and $\nu \in \mathbb{R}_{++}$ such that for all $p \in (0, 1]$ and $\mathbf{x} \in \mathbf{X}_\nu$ satisfying $x_{[\nu]} \leq c$, $(\mathbf{x}, (c)_p) \sim \mathbf{x}$.*

In the case with no risk (i.e., for the subset of allocations with $x_i^p = 1$ for all $i \in \{1, \dots, n(\mathbf{x})\}$), all axioms above are satisfied also by ordinary critical-level utilitarianism. However, as discussed by Arrhenius (2012, Sect. 5.1), critical-level utilitarianism has the properties that adding sufficiently many individuals with wellbeing just above c makes the allocation better than any fixed alternative (thus leading to the Repugnant Conclusion if $c = 0$) and adding sufficiently many individuals with wellbeing just below c makes the allocation worse than any fixed alternative (thus leading to the Very Sadistic Conclusion if $c > 0$). The following axiom ensures that adding individuals at a given level of lifetime wellbeing has bounded importance, thereby avoiding the Repugnant and Very Sadistic Conclusions.

Axiom 7 (Existence of egalitarian equivalence) *For all $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ and $p \in (0, 1]$, if $\mathbf{x} \succ \mathbf{y}$, then there exists $z \in \mathbb{R}$ such that, for all $N \in \mathbb{N}$, $\mathbf{x} \succ (z)_{np} \succ \mathbf{y}$ for some $n \geq N$.*

We will now state our main result, namely that these seven axioms characterize the probability-adjusted rank-discounted critical-level generalized utilitarian SWOs.

Definition 1 An SWR \succsim on \mathbf{X} is a *probability adjusted rank-discounted critical-level generalized utilitarian* (PARDCLU) SWO if there exist $c \in \mathbb{R}_+$, $\delta \in \mathbb{R}_{++}$, and a continuous and increasing function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that, for all $\mathbf{x}, \mathbf{y} \in \mathbf{X}$,

$$\mathbf{x} \succsim \mathbf{y} \Leftrightarrow \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} (u(\mathbf{x}_{[\rho]}) - u(c)) d\rho \geq \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} (u(\mathbf{y}_{[\rho]}) - u(c)) d\rho.$$

Parameter δ is the rank utility discount rate.

Theorem 1 *The following two statements are equivalent.*

- (1) The SWR \succsim satisfies axioms 1–7.
- (2) The SWR \succsim is a PARDCLU SWO.

It follows from the PARDCLU SWO that c is the wellbeing level which, if experienced by an added individual without changing the utilities of the existing population, leads to an alternative which is as good as the original *only if* $\mathbf{x}_{[\nu(\mathbf{x})]} \leq c$. If $\mathbf{x}_{[\nu(\mathbf{x})]} > c$, then there is a context-dependent critical level in the open interval $(c, \mathbf{x}_{[\nu(\mathbf{x})]})$ which depends on the wellbeing levels that exceed c (as well as the probability p with which the added individual exists). This follows from Definition 1, since adding an individual at wellbeing level $\mathbf{x}_{[\nu(\mathbf{x})]}$ increases welfare, while adding an individual at wellbeing level c lowers the weights assigned to individuals at wellbeing levels that exceed c and thereby reduces welfare.

4 Special cases

Cases with no risk correspond to situations where only allocations

$$\mathbf{x} = \left((x_1^w, x_1^p), \dots, (x_{n(\mathbf{x})}^w, x_{n(\mathbf{x})}^p) \right)$$

with $x_i^p = 1$ for all $i = 1, \dots, n(\mathbf{x})$ are considered. In the formulation presented in Section 2, where information arrives in each of T time periods, these correspond to cases where S and, thus, Ω are singleton sets.

The implications of rank-discounted utilitarianism in such settings are discussed in Zuber and Asheim (2012) and Asheim and Zuber (2013). With no risk the modeling here translates exactly to the variable population framework of Asheim and Zuber (2013), while it specializes the fixed population framework of Zuber and Asheim (2012) to a situation with an unbounded but finite number of generations.

Here we highlight special cases *with* risk. First, we show how PARDCLU reduces to rank-dependent expected utilitarianism in the special case where the probability adjusted population size is equal to 1. Second, we discuss to what extent PARDCLU provides a foundation for discounting according to the probability of human extinction, as applied in, e.g., the Stern Review (2007, Ch. 2).

Rank-dependent expected utilitarianism. In the special fixed population case where only allocations \mathbf{x} with probability adjusted population size $\nu(\mathbf{x}) = \sum_{i=1}^{n(\mathbf{x})} x_i^p$ equal to 1 is considered, the result of Theorem 1 leads to *rank-dependent expected utility maximization*—where the decision maker substitutes ‘decision weights’ for probability—but with additional structure. Quiggin (1982) was the first to axiomatize such a theory for decisions under risk, even though the substitution of ‘decision weights’ for probability had been argued by earlier writers to explain behavior inconsistent with the vNM theory.

In the formulation presented in Section 2, where information arrives in each of T time periods, this corresponds to the case where $T = 1$, so that $\Omega = S$, and where $n^1(s) = 1$ for all $s \in S$, so that one individual lives independently of how the risk is resolved. Even though we thereby depart from our basic setting without individual risk, we may interpret this as one person being subject to a lottery where the prizes $(w_1^1(s_1), \dots, w_1^1(s_{|S|}))$ are won with probabilities $(p^1(s_1), \dots, p^1(s_{|S|}))$.

Let $s_{[]} = (s_{[1]}, \dots, s_{[|S|]})$ denote a reordering of s that turns $(w_1^1(s_1), \dots, w_1^1(s_{|S|}))$ into a non-decreasing profile: $w_1^1(s_{[r]}) \leq w_1^1(s_{[r+1]})$ for all ranks $r = 1, \dots, |S| - 1$. Then PARDCLU implies preferences for lotteries that are represented by:

$$\sum_{r=1}^{|S|} h_r(p) u(w_1^1(s_{[r]})),$$

where the probability weighting functions $h_r : \Delta(S) \rightarrow [0, 1]$ are defined by

$$h_r(p) = f\left(\sum_{r'=1}^r p^1(s_{[r']})\right) - f\left(\sum_{r'=1}^{r-1} p^1(s_{[r']})\right),$$

with $f : [0, 1] \rightarrow [0, 1]$ given by $f(\rho) = (1 - e^{-\delta\rho}) / (1 - e^{-\delta})$ and using the convention $\sum_{r'=1}^0 p^1(s_{[r']}) = 0$.² Note that the function f is concave; the plausibility of this property is discussed by Quiggin (1987). In addition, our axioms (in particular, Axiom 5) lead to the special exponential structure implied by the function f . As can be easily checked by applying l'Hôpital's rule, f approaches the identity function

²This follows from Definition 1 by integrating the utility weights $e^{-\delta\rho}$, leading to the following cumulative utility weights: $\int_0^\rho e^{-\delta\rho'} d\rho' = -(e^{-\delta\rho} - 1) / \delta$. The function f is determined by multiplying these cumulative weights by $\delta / (1 - e^{-\delta})$ so that $f(0) = 0$ and $f(1) = 1$.

as $\delta \downarrow 0$. Thus, if the probability adjusted population size equals 1, then PARDCLU approaches ordinary expected utility maximization as rank-discounting vanishes.

Human extinction. By appealing to Harsanyi's (1953) original position and using Harsanyi's (1955) theorem, Dasgupta and Heal (1979, pp. 269–275) justified the use of discounted utilitarianism where the utility discount rate is the probability of human extinction. Also the Stern Review (2007, Ch. 2) argued that this probability is the primary justification for utility discounting (other contributions include Bommier and Zuber, 2008, and Roemer, 2011). Blackorby, Bossert and Donaldson (2007) supported this justification within a variable population framework. To what extent is PARDCLU consistent with this position?

The variable population case where population remains constant up to the time of human extinction can be captured in the formulation presented in Section 2, where information arrives in each of T time periods, by having $T > 1$, letting the set of potential signals, S , equal $\{s_c, s_e\}$, where s_c signals continued existence and s_e signals extinction, and assuming that population is constant (and normalized to 1) up to the first time t at which the signal at t , s^t , equals s_e :

$$n^t(s^1, \dots, s^t) = \begin{cases} 1 & \text{if } s^\tau = s_c \text{ for all } \tau = 1, \dots, t, \\ 0 & \text{otherwise.} \end{cases}$$

Normalizing population to 1 amounts to assuming no *intragenerational* inequality.

Assume that s^t is i.i.d. where s_c is observed with probability π and s_e is observed with probability $1-\pi$, implying that the probability of continued existence in period t is π^t : $p^t(s_c, \dots, s_c) = \pi^t$. If well-being is correlated with time so that $w_1^t(s_c, \dots, s_c) \leq w_1^{t+1}(s_c, \dots, s_c)$ for all times $t = 1, \dots, T-1$, then PARDCLU implies preferences over streams that are represented by:

$$\sum_{t=1}^T \left[f\left(\frac{\pi(1-\pi^t)}{1-\pi}\right) - f\left(\frac{\pi(1-\pi^{t-1})}{1-\pi}\right) \right] u(w_1^t(s_c, \dots, s_c)),$$

where, as above, $f: \mathbb{R}_+ \rightarrow \mathbb{R}_+$ is given by $f(\rho) = (1 - e^{-\delta\rho})/(1 - e^{-\delta})$, but with an extended domain. This follows from Definition 1 and the argument of footnote 2 by noting that $\pi + \dots + \pi^t = \pi(1 - \pi^t)/(1 - \pi)$.

Note that as $\delta \downarrow 0$, f approaches the identity function:

$$f\left(\frac{\pi(1-\pi^t)}{1-\pi}\right) - f\left(\frac{\pi(1-\pi^{t-1})}{1-\pi}\right) \rightarrow \frac{\pi}{1-\pi}(\pi^{t-1} - \pi^t) = \pi^t.$$

Therefore, as rank-discounting vanishes, PARDCLU approaches the principle of discounting utility according to the probability of human extinction, as applied by the Stern Review (2007, Ch. 2). However, for $\delta > 0$, PARDCLU implies that *utility* is discounted according to both rank and the probability of human extinction. If well-being is correlated with time—which is the case considered above—discounting according to rank and the probability of human extinction reinforce each other, while they might pull in opposite directions otherwise. In all cases, *wellbeing* is also discounted according to the absolute wellbeing level if the function u is strictly concave, so that wellbeing is transformed into utility at a decreasing rate.

5 The length of the planning horizon

Despite the normative appeal of the PARDCLU approach, as expressed by the axioms characterizing it, it raises several issues that need to be discussed. We do so in this and the two subsequent sections.

One important feature of evaluation based on PARDCLU is its dependence on the planning horizon and the risk of the planning horizon. It is important to note that, according to PARDCLU, it is the total population, rather than the planning horizon, that matters. In particular, social evaluation based on PARDCLU is completely indifferent between having 10 billion people alive for 100 years and 1 billion people alive for 1000 years if all have the same wellbeing and live for sure, as total population is the same in both alternatives. One may object to this conclusion on the basis that people might prefer to live in a society with more people (so as to have richer scope for social interactions), or on the contrary to have more descendants.

Two responses can be made to this objection. First, one can argue that if these sentiments are valid components of welfare, they should be included in the x_i^w numbers, so that the equality of these numbers in the two situations means

that these phenomena are appropriately taken into account. Second, one might argue that such interpersonal welfare effects are not part of a well-defined theory of wellbeing so that a theory of justice should not take them into account.

It must be noted that this feature actually extends to the case where population size is risky. If wellbeing is perfectly equal, then only expected total population size matters, so that society is completely risk neutral with respect to the risk of population size. Evaluation based on PARDCLU is indifferent whether n people exist for sure, or n_1 people exist with probability p and n_2 people with probability $1 - p$, provided that $pn_1 + (1 - p)n_2 = n$. This is in stark contrast with criteria exhibiting catastrophe avoidance in the sense of Bommier and Zuber (2008).

A more practical problem has to do with the possible discrepancy between the actual length of society and the planning horizon considered in applied models. It is a common practice in applied model, for instance in integrated climate-economy assessment models, to assume that the economy reaches a steady state at a certain point (the planning horizon), and to either neglect what happens next, or use a simple recursive formula to value the future. In particular, the actual length of the society is not taken into account. However, it is not possible to neglect the far future in such a way in the PARDCLU approach.

In the following we present a method for dealing with this issue. Assume a dynastic model with n dynasties, where there is risk about the wellbeing of the dynasties. All dynasties reach a steady state wellbeing level simultaneously, with the steady state distribution of welfare and probability of existence being given by $\bar{\mathbf{x}} = ((\bar{x}_1^w, \bar{x}_1^p), \dots, (\bar{x}_n^w, \bar{x}_n^p))$. Let \mathbf{y} the distribution of wellbeing and probability before attaining the limiting distribution.

Assume that the duration of all dynasties is subject to the same risk. Conditional on dynasty i existing when we reach the steady state (which occurs with probability \bar{x}_i^p) the dynasty will last for exactly T more generations with probability π_T . Hence, conditional on dynasty i existing, generation T of the dynasty will exist with probability $q_T = \sum_{t \geq T} \pi_t$ (where generation 0 is the label for the generation

existing when the steady state is reached). The expected number of generations after reaching the steady state is $N = \sum_{t \geq 0} \pi_t t = \sum_{t \geq 0} q_t$ (assumed to be finite).

In such a framework the actual allocation we should consider is

$$\mathbf{x} = (\mathbf{y}, (\bar{x}_1^w, \bar{x}_1^p), \dots, (\bar{x}_1^w, q_1 \bar{x}_1^p), \dots, \dots, (\bar{x}_n^w, \bar{x}_n^p), \dots, (\bar{x}_n^w, q_n \bar{x}_n^p), \dots).$$

In the steady state, all individuals of dynasty i have the same welfare level \bar{x}_i^w and they have probabilities of existence $\bar{x}_i^p, q_1 \bar{x}_i^p, \dots, q_t \bar{x}_i^p, \dots$. Hence the probability-adjusted number of people at welfare level \bar{x}_i^w is $\sum_t q_t \bar{x}_i^p = N \bar{x}_i^p$. Hence, using Axiom 3, we can rewrite \mathbf{x} as follows: $\mathbf{x} = (\mathbf{y}, (\bar{x}_1^w, N \bar{x}_1^p), \dots, (\bar{x}_n^w, N \bar{x}_n^p))$.

If we had neglected the generations in the far future, we would instead have considered the allocation $\mathbf{z} = (\mathbf{y}, (\bar{x}_1^w, \bar{x}_1^p), \dots, (\bar{x}_n^w, \bar{x}_n^p))$. Obviously, by using such a naive approach where the far future is neglected, the welfare weights on the (transformed) utilities in the PARDCLU formula are changed. More precisely, assuming that $\bar{x}_i^w < \bar{x}_{i+1}^w$, the weight in the naïve approach on the transformed utility $u(x_k^w)$ of people with welfare x_k^w between \bar{x}_i^w and \bar{x}_{i+1}^w must be multiplied by a factor $e^{-\delta(N-1) \sum_{j \leq i} \bar{x}_j^p} (< 1)$ to arrive at the weights in the actual PARDCLU formula.

In the limit where $N \rightarrow +\infty$, the weights on the transformed utility of people with welfare higher than $\bar{x}_1^w = \min_{i \in \{1, \dots, n\}} \bar{x}_i^w$ would be negligible. If function u is bounded, this might mean that anything that happens to people with welfare higher than the lowest welfare of the limiting distribution is unimportant from the social choice point of view. This gives a clear indication where we should provide welfare improvements.

Although we have argued that neglecting the future when there is a limiting distribution can be quite misleading, the above discussion also provides a fix that can be used in practical applications of the class of PARDCLU criteria. If we know the expected number of future generations N , then we can easily construct the allocation \mathbf{x} using the information on the limiting distribution $\bar{\mathbf{x}}$, by simply changing the probability of the existence of last generations from x_i^p to $N x_i^p$.

6 Time consistency

The decision maker will be time inconsistent when decisions in later periods do not coincide with the plan that was made in earlier periods. Hammond (1983) suggested that if social decision-making is consequentialist (that is, if social situations in each state of the world are assessed only on basis of their consequences in this state of the world), time consistency is ensured only by using an expected utility criterion.

The PARDCLU is not an expected utility. Hence, if decision making is consequentialist, PARDCLU are bound to yield time inconsistencies. Note that the issue also arises when there is no risk, as discussed in Asheim and Zuber (2013). Indeed, rank discount factors may depend on the relative position of past generations so that, when the PARDCLU is used at later periods ignoring the past, it may chose a different plan than decided earlier, because the relative weight on the wellbeing of future generations has changed.³ The problem of time consistency when there is risk is more severe because not only the past, but also unrealized states of the world matter. This dependence on unrealized states of the world is not specific to the PARDCLU approach, for it also arises for criteria such as those suggested by Diamond (1967), Epstein and Segal (1992) and Grant et al. (2010).

Given the result by Hammond (1983), two options are possible: reject consequentialism or time consistency. If one wants to avoid the time inconsistency problem, one has to abandon the hypothesis of time invariant and consequentialist decision making. This involves defining social preferences that are conditional on the past and on unrealized events. Consider again the dynamic framework introduced in Section 2, where the allocations \mathbf{x} and \mathbf{y} are derived from probability distributions p and q , for fixed functions n^t and w^t determining the numbers of individuals and the distribution of wellbeing at every time t . Also assume that there is a history of

³However, as shown by Zuber and Asheim (2012, Section 6) in the setting of the Ramsey and Dasgupta-Heal-Solow models of optimal growth, it might happen that no time inconsistency arise at the optimal plan for a rank-discounted utilitarian criterion. This property is not be true in general and depends on the precise economic model which is considered.

signal $\omega^t = \{s^1, \dots, s^t\}$ such that

- (1) $p^\tau(\tilde{\omega}^\tau) = q^\tau(\tilde{\omega}^\tau)$ for all $\tilde{\omega} \in \Omega$ and $\tau < t$, and
- (2) $p^\tau(\tilde{\omega}^\tau) = q^\tau(\tilde{\omega}^\tau)$ for all $\tau \geq t$ and all $\tilde{\omega} = (\tilde{s}^1, \dots, \tilde{s}^t, \tilde{s}^{t+1}, \dots, \tilde{s}^T) \in \Omega$ such that $(\tilde{s}^1, \dots, \tilde{s}^t) \neq (s^1, \dots, s^t)$.

This means that \mathbf{x} and \mathbf{y} are allocations where people in the past and unrealized states of the world have the same welfare and probability of existence. We can define \mathbf{x}_{ω^t} and \mathbf{y}_{ω^t} as the allocations for people in future generations and still possible states of the world.⁴ Then we can simply formulate the social preference ordering conditional on the history ω^t , denoted \succeq_{ω^t} in the following way:

$$\mathbf{x}_{\omega^t} \succeq_{\omega^t} \mathbf{y}_{\omega^t} \iff \mathbf{x} \succeq \mathbf{y}.$$

Such a construction would trivially imply time consistency.

Such a construction is actually the solution proposed by Epstein and Segal (1992) when they develop a formula to update the weights on individuals' utilities ex post to obtain consistent planning. What the updating rule does is to aggregate the information needed on unrealized alternatives, so that ex post decision making does depend on unrealized alternatives in a proper way. The issue is rather that the updating rule for social preferences may be rather complex. This solution may be possible for some problems when the amount of information required is not too important. But for most dynamic decision problems involving many unrealized alternatives this may be difficult to implement.

The second option would be to accept the possibility of time inconsistencies in social decision making. There are still situations where time consistency problem does not arise for PARDCLU, even if we stick to consequentialist decision making. Axioms 4 and 5 precisely describe situations where the information on the past or

⁴The definition of \mathbf{x}_{ω^t} would be

$$\mathbf{x}_{\omega^t} = (((w_j^\tau(\omega^\tau), p^\tau(\omega^\tau))_{j \in \{1, \dots, n^\tau(\omega^t)\}})_{\omega^\tau \in \{s^1\} \times \dots \times \{s^t\} \times S^{\tau-t}})_{\tau \in \{t, \dots, T\}},$$

and similarly for \mathbf{y}_{ω^t} by substituting $q^\tau(\omega^\tau)$ by $p^\tau(\omega^\tau)$.

unrealized alternatives is not needed to evaluate allocations. It should be clear that only very specific problems would permit to use the independence Axioms 4 and 5: as soon as someone in the past (or in unrealized states of the world) has a welfare level in between the welfare levels of two future (potential) individuals he will matter for decision making.

Then, if one anticipates that the decision process will be time invariant and consequentialist, the only way to mitigate the time inconsistency problem is to devise a sophisticated planning strategy (see Pollak, 1968; Blackorby et al., 1973, for early references). The main idea is that we can anticipate today that the time consistency problem will arise tomorrow, and therefore chose a policy that will induce the next generations to do the best choices according to the use of an PARDCLU criterion today. Typically, the sophisticated planning solution does not yield a solution which is as good as the optimal plan if we could commit to a policy at any future decision node. But it may improve upon a naive solution which does not anticipate the problem, and it yields consistent planning. The exact form of the sophisticated planning solution, and how far it can go in reducing the suboptimality will depend on the specific economic problem under consideration. But it seems to be a promising route when one use criteria such as PARDCLU that do not satisfy strong enough separability and time invariance properties.

7 Ex ante vs ex post approaches

A seminal result in the literature on social choice in risky situations is Harsanyi's theorem (Harsanyi, 1955). Harsanyi proved that, in the context of risk, social rationality (embodied in the expected utility assumption) and the Pareto principle impose severe constraints on the form of the social welfare function. Specifically, the social criterion should be a linear combination of individuals' expected utility, ruling out preference for (utility) redistribution both ex ante and ex post.

Since then, the literature has hesitated between an ex ante approach that relaxes rationality (Diamond, 1967; Epstein and Segal, 1992; Grant et al., 2010) to allow for

ex ante fairness, and an *ex post* approach that fails the Pareto principle (Broome, 1991; Hammond, 1983; Fleurbaey, 2010) to allow for *ex post* fairness.

How does the PARDCLU approach fit in this debate? This is not clear, because we interpret individuals in different states of the world as essentially different individuals. The key issue in the *ex ante* versus *ex post* debate is whether we should respect individuals' *ex ante* preference, which is not possible to formulate in our framework because we do not assume that individuals face risk: only the society faces some risk on the identity of individuals and welfare distribution. One possibility though would be to interpret the utility numbers x_i^w as incorporating the risk actual individuals face during their lifetime. In that case, and if the welfare index is concordant with individuals' *ex ante* preferences, one could consider the PARDCLU approach as an *ex ante* approach in the sense that it respects individuals' risk preferences. It would therefore be related to criteria suggested by Diamond (1967), Epstein and Segal (1992) and Grant et al. (2010). As we discussed in Section 6, PARDCLU criteria actually face the same difficulty in terms of time inconsistency as these *ex ante* criteria.

This time inconsistency is related to the fact that the PARDCLU are not *ex post* criteria, in a sense somewhat different from the one discussed in the *ex ante* versus *ex post* debate. Following Fleurbaey (2010), one could define an *ex post* approach as one first assessing social welfare in each state of the world, and then performing an aggregation of these *ex post* welfare judgements to assess a risky situation. There is clear tension between such an approach and our axiom 3 (probability-adjusted Suppes-Sen). Assume like Fleurbaey (2010) that the social criterion is the expected value of an equally-distributed equivalent function. There are two equiprobable states of the world and two individuals in each state of the world. Consider the following allocations: in allocation \mathbf{x} the distribution of welfare is (\underline{z}, \bar{z}) in each state of the world; in allocation \mathbf{y} , the distribution of welfare is $(\underline{z}, \underline{z})$ in one state of the world and (\bar{z}, \bar{z}) in the other state of the world. The probability-adjusted Suppes-Sen axiom imply that \mathbf{x} and \mathbf{y} should be deemed equivalent. However, if

social welfare is inequality-averse (with respect to the distribution of welfare), the equally-distributed equivalent in each state of the world in \mathbf{x} should be less than $(\underline{z} + \bar{z})/2$ which is the expected social welfare in \mathbf{y} . Inequality averse expected equally-distributed equivalent criteria would therefore prefer \mathbf{y} to \mathbf{x} , contradicting the probability-adjusted Suppes-Sen axiom.

The *ex post* approach relies on a consequentialist assumption that the welfare assessment of situations in each state of the world should only depend on the consequences in this particular state of world. One might want to question this consequentialist assumption. For instance, should the ex ante relative priority of two individuals located in the same state of the world be independent of what occurs in different states of the world? Or should the situation of people in other states of the world modify the way we view redistribution in a particular state of the world. In particular, we may well accept more inequality in good states when there is some catastrophic state, because the real priority is to raise the welfare of people in that state, which is something we would not do otherwise. This is the kind of intuition that may justify our PARDCLU approach.

If one does not accept the non-consequentialist intuition, a possibility could be to follow the expected equally-distributed equivalent approach *à la* Fleurbaey (2010) and take the expected value of a RDCLU social welfare function. This is a direction we do not investigate in the present paper.

8 Concluding remarks

In conjunction with our companion paper, Asheim and Zuber (2013), the present paper contributes to the fields of population ethics and social evaluation in risky situations by proposing and axiomatizing the probability adjusted rank-discounted critical-level generalized utilitarian (PARDCLU) SWO. By doing so we have taken an important step towards preparing the rank-discounted utilitarian (RDU) criterion (see Zuber and Asheim, 2012) for practical use, e.g. for evaluation of climate policies and other policy issues with long-run consequences.

We have shown how the PARDCLU SWO reduces to rank-dependent expected utility maximization with additional structure in the special case where the probability adjusted population size equals 1, and can be used to handle the situation where there is a positive probability of human extinction.

When evaluating consequences that stretch centuries into the future, it seems less important to consider the fluctuations in wellbeing and individual risk that people face during their own lifetimes. Rather, the important issues are interpersonal inequality and the social risk associated with what level of wellbeing future people will experience in the world they will be born into.

Consequently, we have presented a framework where individuals live for one period only and are not subject to individual risk. In this framework one cannot differentiate between inequality aversion, fluctuation aversion, and risk aversion—a distinction that is sometimes highlighted in literature on climate change evaluation—only inequality aversion matters in the present context.

In Zuber and Asheim (2012, Section 6), where it is imposed through the structure that there are infinitely many time periods, we show how RDU leads to sustainable outcomes in models of economic growth. This basic support for sustainability does not extend to the present criterion with endogenous population size and probability of existence, where the main concern is to avoid lives with low wellbeing. A stark conclusion is that it might be social preferable to increase the per-period probability of extinction if per capita wellbeing is decreasing over time, as this increases the utility weight on the better-off earlier generations.

In general, there might be an argument in favor of distinguishing the conception of justice from the forces (like altruism) that are instrumental in attaining it, e.g., if impartiality follows from considering an original position where individuals do not have extensive times of natural sentiments (Rawls, 1971, p. 129). However, considering the social context in which people live seems essential when applying the PARDCLU SWO in a setting where population size and probability of existence are endogenous. In particular, a more pro-natal implication would follow if we

assume that the wellbeing of individuals depends also on their reproductive choices, so that wellbeing of one generation increases with the size and living conditions of the next.

References

- Arrhenius, G. (2000). *Future Generations – A Challenge for Moral Theory*. PhD dissertation, Uppsala University.
- Arrhenius, G. (2012). *Population Ethics – The Challenge of Future Generations*. Forthcoming.
- Asheim, G.B., Zuber, S., (2013). “Escaping the repugnant conclusion: rank-discounted utilitarianism with variable population”, Memorandum 23/2012, Department of Economics, University of Oslo. Revised April 2013. Forthcoming in *Theoretical Economics*.
- Blackorby, C., Bossert, W., Donaldson, D. (2007). Variable-population extensions of social aggregation theorems. *Social Choice and Welfare* **28**, 567–589.
- Blackorby, C., Nissen, D., Primont, D., Russell, R.R (1973). Consistent intertemporal decision making. *Review of Economic Studies* **40**, 239–248.
- Bommier, A., Zuber, S. (2008). Can preferences for catastrophe avoidance reconcile social discounting with intergenerational equity? *Social Choice and Welfare* **31**, 415–434.
- Broome, J. (1991). *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Blackwell.
- Dasgupta, P.S., Heal, G.M. (1979). *Economic Theory and Exhaustible Resources*. Cambridge, UK: Cambridge University Press.
- Diamond, P. (1967). Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: Comment. *Journal of Political Economy* **75**, 765–766.
- Epstein, L.G., Segal, U. (1992). Quadratic social welfare functions. *Journal of Political Economy* **100**, 691–712.
- Fleurbaey, M. (2010). Assessing risky social situations. *Journal of Political Economy* **118**, 649–680.
- Grant, S., Kajii, A., Polak, B., Safra, Z. (2010). Generalized Utilitarianism and Harsanyi’s Impartial Observer Theorem. *Econometrica*, **78**, 1939–1971.

- Hammond, P. (1983). Ex post optimality as a dynamically consistent objective for collective choice under uncertainty. In *Social Choice and Welfare*, eds. P. Pattanaik and M. Salles, pp. 175–205 Amsterdam: North-Holland.
- Harsanyi, J.C. (1953). Cardinal utility in welfare economics and in the theory of risk-taking. *Journal of Political Economy* **61**, 434–435.
- Harsanyi, J.C. (1955). Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of Political Economy* **63**, 309–321.
- Koopmans, T.C. (1960). Stationary ordinal utility and impatience, *Econometrica* **28**, 287–309.
- Parfit, D. (1976). “On doing the best for our children”. In *Ethics and Population*, ed. M. Bayles, pp. 100–102 Cambridge: Schenkman.
- Parfit, D. (1982). Future generations, further problems. *Philosophy and Public Affairs* **11**, 113–172.
- Parfit, D. (1984). *Reasons and Persons*. Oxford: Oxford University Press.
- Pollak, R.A. (1968). Consistent planning. *Review of Economic Studies* **35**, 201–208.
- Quiggin, J. (1982), “A theory of anticipated utility”, *Journal of Economic Behavior and Organization* **3**, 323–43.
- Quiggin, J. (1987), “Decision weights in anticipated utility theory”, *Journal of Economic Behavior and Organization* **8**, 641–645.
- Rawls, J. (1971). *A Theory of Justice*, 1999 revised edition, Cambridge (Mass.): The Belknap Press of the Harvard University Press.
- Roemer, J.E. (2011). The ethics of intertemporal distribution in a warming planet. *Environmental and Resource Economics* **48**, 363–390.
- Stern, N. (2007). *The Stern Review on the Economics of Climate Change*. HM Treasury, London.
- Zuber, S., Asheim, G.B. (2012). “Justifying social discounting: the rank-discounted utilitarian approach”, *Journal of Economic Theory* **147**, 1572–1601.

A Proof of Theorem 1

To prove the Theorem 1, we need to introduce subsets of \mathbf{X} . For any $k \in \mathbb{N}$, denote by $\mathbf{X}_{1/k} = \{\mathbf{x} \in \mathbf{X} : x_i^p = 1/k, \forall i \in \{1, \dots, n(\mathbf{x})\}\}$ the set of allocations where all individuals have the same rational probability $1/k$ of existing. Denote by $\mathbf{X}_{\mathbb{Q}_{++}} = \{\mathbf{x} \in \mathbf{X} : x_i^p \in \mathbb{Q}_{++}, \forall i \in \{1, \dots, n(\mathbf{x})\}\}$ the set of allocations where all individuals have probabilities of existing which are positive rational numbers.

It is straightforward to show that (2) implies (1) in Theorem 1. We show that (1) implies (2) by proving the four following lemmas.

We start with Lemmas 1 and 2 which establish how the representation result of Asheim and Zuber (2013) can be extended to the present case as long as the probabilities of existence are given by rational numbers.

Lemma 1 *If Axioms 1–7 hold, then there exists $c \in \mathbb{R}_+$, $\delta \in \mathbb{R}_{++}$ and a continuous and increasing function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that for any $k \in \mathbb{N}$ for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_{1/k}$,*

$$\mathbf{x} \succsim \mathbf{y} \iff \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} (u(\mathbf{x}_{[\rho]}) - u(c)) d\rho \geq \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} (u(\mathbf{y}_{[\rho]}) - u(c)) d\rho$$

Proof. For any $k \in \mathbb{N}$, Axioms 1–7 above restricted to $\mathbf{X}_{1/k}$ collapse to Axioms 1–7 of Asheim and Zuber (2013), provided we take $p = 1/k$ in Axioms 4–7. Hence, by Theorem 1 of Asheim and Zuber (2013) there exist $\beta_{1/k} \in (0, 1)$ and a continuous increasing function $u_{1/k} : \mathbb{R} \rightarrow \mathbb{R}$ such that, for all $\mathbf{x}, \mathbf{y} \in \mathbf{X}_{1/k}$, $\mathbf{x} \succsim \mathbf{y}$ if and only if

$$\begin{aligned} (1 - \beta_{1/k}) \sum_{r=1}^{n(\mathbf{x})} \beta_{1/k}^{r-1} (u_{1/k}(x_{\pi(r)}^w) - u_{1/k}(c)) \\ \geq (1 - \beta_{1/k}) \sum_{r=1}^{n(\mathbf{y})} \beta_{1/k}^{r-1} (u_{1/k}(y_{\pi(r)}^w) - u_{1/k}(c)), \end{aligned} \quad (\text{A1})$$

where the critical level parameter c is determined by Axiom 6 and is therefore independent of k , and where the factor $1 - \beta_{1/k}$ ensures that utility weights sum up to $1 - \beta_{1/k}^{n(\mathbf{x})}$ and $1 - \beta_{1/k}^{n(\mathbf{y})}$ respectively.

Consider any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_1$ such that $n(\mathbf{x}) = n(\mathbf{y}) = n$. For any $k \in \mathbb{N}$, construct $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{1/k}$ such that $n(\hat{\mathbf{x}}) = n(\hat{\mathbf{y}}) = nk$ and, for any $i \in \{1, \dots, n\}$, $\hat{x}_{ki-j}^w = x_i^w$ and $\hat{y}_{ki-j}^w = y_i^w$ for all $j \in \{0, \dots, k-1\}$. By construction, $\nu(\mathbf{x}) = \nu(\mathbf{y}) = \nu(\hat{\mathbf{x}}) = \nu(\hat{\mathbf{y}}) = n$, $\mathbf{x}_{[\cdot]} = \hat{\mathbf{x}}_{[\cdot]}$ and $\mathbf{y}_{[\cdot]} = \hat{\mathbf{y}}_{[\cdot]}$. By Axioms 1, 2 and 3, we have $\mathbf{x} \succsim \mathbf{y} \iff \hat{\mathbf{x}} \succsim \hat{\mathbf{y}}$, and therefore, using the above representation:

$$(1 - \beta_1) \sum_{r=1}^n \beta_1^{r-1} u_1(\mathbf{x}(\pi(r))) \geq (1 - \beta_1) \sum_{r=1}^n \beta_1^{r-1} u_1(\mathbf{y}(\pi(r)))$$

$$\begin{aligned}
&\iff (1 - \beta_{1/k}) \sum_{r'=1}^{nk} \beta_{1/k}^{r'-1} u_{1/k}(\hat{\mathbf{x}}(\pi(r'))) \geq (1 - \beta_{1/k}) \sum_{r'=1}^{nk} \beta_{1/k}^{r'-1} u_{1/k}(\hat{\mathbf{y}}(\pi(r'))) \\
&\iff (1 - \beta_{1/k}^k) \sum_{r=1}^n (\beta_{1/k}^k)^{r-1} u_{1/k}(\mathbf{x}(\pi(r))) \\
&\qquad\qquad\qquad \geq (1 - \beta_{1/k}^k) \sum_{r=1}^n (\beta_{1/k}^k)^{r-1} u_{1/k}(\mathbf{y}(\pi(r)))
\end{aligned}$$

since $(1 - \beta_{1/k}) \sum_{r'=1}^k \beta_{1/k}^{r'-1} = 1 - \beta_{1/k}^k$. Because additive representations are unique up to an affine transformation, the above equivalence implies $\beta_{1/k}^k = \beta_1$ and that we can set $u_{1/k} = u_1$, using the normalization $u_{1/k}(0) = u_1(0) = 0$.

Denoting $p = 1/k$ and $\delta = -\ln \beta_1$, this implies that $\beta_{1/k} = (\beta_1)^{1/k} = e^{-\delta p}$. Moreover, since

$$(1 - \beta_p) \beta_p^{r-1} = (1 - e^{-\delta p}) e^{-\delta p(r-1)} = e^{-\delta p(r-1)} - e^{-\delta pr} = \delta \int_{p(r-1)}^{pr} e^{-\delta \rho} d\rho,$$

and by denoting $u = u_1$ we can rewrite inequality (A1) as:

$$\delta \int_0^{\nu(\mathbf{x})} e^{-\delta \rho} (u(\mathbf{x}_{[\rho]}) - u(c)) d\rho \geq \delta \int_0^{\nu(\mathbf{y})} e^{-\delta \rho} (u(\mathbf{y}_{[\rho]}) - u(c)) d\rho,$$

where $\nu(\mathbf{x}) = n(\mathbf{x})/k$ and $\nu(\mathbf{y}) = n(\mathbf{y})/k$. This establishes Lemma 1. ■

Lemmas 2 is concerned only with the same-number case, as a separate argument has to be used anyway for the case where the difference between the probability-adjusted population size of two different allocations is irrational.

Lemma 2 *If Axioms 1–7 hold, then there exists $c \in \mathbb{R}_+$, $\delta \in \mathbb{R}_{++}$ and a continuous and increasing function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that for any $\nu \in \mathbb{Q}_{++}$, for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_{\mathbb{Q}_{++}}$ such that $\nu(\mathbf{x}) = \nu(\mathbf{y}) = \nu$,*

$$\mathbf{x} \succsim \mathbf{y} \iff \int_0^\nu e^{-\delta \rho} u(\mathbf{x}_{[\rho]}) d\rho \geq \int_0^\nu e^{-\delta \rho} u(\mathbf{y}_{[\rho]}) d\rho$$

Proof. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_{\mathbb{Q}_{++}}$ such that $\nu(\mathbf{x}) = \nu(\mathbf{y}) = \nu$, let k be the least common denominator of all the probabilities in the two allocations. This means that for all $i \in \{1, \dots, n(\mathbf{x})\}$ there exists a positive integer $\ell_i^{\mathbf{x}}$ such that $x_i^p = \ell_i^{\mathbf{x}}/k$. Similarly, for all $i \in \{1, \dots, n(\mathbf{y})\}$, there exists a positive integer $\ell_i^{\mathbf{y}}$ such that $y_i^p = \ell_i^{\mathbf{y}}/k$.

We can construct $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{1/k}$ in the following way:⁵

- (a) For any $i \in \{1, \dots, n(\mathbf{x})\}$, $\hat{x}_{\ell + \sum_{j=1}^{i-1} \ell_j^{\mathbf{x}}}^w = x_i^w$, for all $\ell \in \{1, \dots, \ell_i^{\mathbf{x}}\}$;
- (b) For any $i \in \{1, \dots, n(\mathbf{y})\}$, $\hat{y}_{\ell + \sum_{j=1}^{i-1} \ell_j^{\mathbf{y}}}^w = y_i^w$, for all $\ell \in \{1, \dots, \ell_i^{\mathbf{y}}\}$.

⁵Using the convention $\sum_{j=1}^0 \ell_j^{\mathbf{x}} = \sum_{j=1}^0 \ell_j^{\mathbf{y}} = 0$.

By construction, $\nu(\mathbf{x}) = \nu(\mathbf{y}) = \nu(\hat{\mathbf{x}}) = \nu(\hat{\mathbf{y}}) = n$, $\mathbf{x}_{[1]} = \hat{\mathbf{x}}_{[1]}$ and $\mathbf{y}_{[1]} = \hat{\mathbf{y}}_{[1]}$, and

$$\begin{aligned} \mathbf{x} \succsim \mathbf{y} &\iff \hat{\mathbf{x}} \succsim \hat{\mathbf{y}} \\ &\iff \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho \geq \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho \end{aligned}$$

by Axioms 1, 2, 3, and Lemma 1. ■

Lemma 3 shows how the representation of Lemma 2 in the same-number case (where the compared allocations have the same probability-adjusted population size) can be applied also when probabilities of existence are allowed to be irrational, using the property that the rational numbers are dense in the real numbers.

Lemma 3 *If Axioms 1–7 hold, then there exist $c \in \mathbb{R}_+$, $\delta \in \mathbb{R}_{++}$ and a continuous and increasing function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that for any $\nu \in \mathbb{R}_{++}$, for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_\nu$,*

$$\mathbf{x} \succsim \mathbf{y} \iff \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho \geq \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho \quad (\text{A2})$$

Proof. Consider any $\nu \in \mathbb{R}_{++}$, and any $\mathbf{x}, \mathbf{y} \in \mathbf{X}_\nu$. If $\mathbf{x}, \mathbf{y} \in \mathbf{Q}_{++}$ (so that $\nu \in \mathbb{Q}_{++}$), then Lemma 2 yields the result. Assume therefore that $\mathbf{x}, \mathbf{y} \notin \mathbf{Q}_{++}$ and more specifically that $x_i^p \in \mathbf{Q}_{++}$ for all $i \in \{1, \dots, n(\mathbf{x}) - 1\}$, $y_j^p \in \mathbf{Q}_{++}$ for all $j \in \{1, \dots, n(\mathbf{y}) - 1\}$, and $x_{n(\mathbf{x})}, y_{n(\mathbf{y})} \notin \mathbf{Q}_{++}$. Assuming that the last individual is the one with an irrational probability of existing is made without loss of generality because of Axiom 3. Extension of the proof to more than one individual with an irrational probability of existing is similar to the one developed below. Because of Axiom 1, equivalence (A2) holds if and only if the following equivalence holds:

$$\mathbf{x} \succ \mathbf{y} \iff \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho > \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho. \quad (\text{A3})$$

$$\text{Step 1: } \mathbf{x} \succ \mathbf{y} \implies \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho > \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho.$$

Assume that $\mathbf{x} \succ \mathbf{y}$. By Axiom 2, there exists $\tilde{\mathbf{x}} \in \mathbf{X}_\nu$ such that $\tilde{\mathbf{x}}_{[1]} < \mathbf{x}_{[1]}$ and $\tilde{\mathbf{x}} \succ \mathbf{y}$. It is sufficient to show $\int_0^\nu e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho \geq \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho$ since then it follows by the definitions of the step functions $\mathbf{x}_{[1]}$ and $\tilde{\mathbf{x}}_{[1]}$ that $\mathbf{x}_{[1]} > \tilde{\mathbf{x}}_{[1]}$ implies $\int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho > \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho$.

Let $\hat{\nu} \in \mathbf{Q}_{++}$ such that $0 < \hat{\nu} - \nu < 1$, and denote $\hat{p} = \hat{\nu} - \nu$. Let $p_{\tilde{x}} \in \mathbf{Q}_{++}$ be such that $0 < p_{\tilde{x}} < \tilde{x}_{n(\tilde{\mathbf{x}})}^p$ and denote $\varepsilon_{\tilde{x}} = \tilde{x}_{n(\tilde{\mathbf{x}})}^p - p_{\tilde{x}}$. Likewise, let $p_y \in \mathbf{Q}_{++}$ be such that $x_{n(\mathbf{y})}^p < p_y < \hat{p}$ and denote $\varepsilon_y = p_y - x_{n(\mathbf{y})}^p$.

Let $z = \max\{\tilde{\mathbf{x}}_{[\nu]}, \mathbf{y}_{[\nu]}\}$. By Axiom 4,

$$\tilde{\mathbf{x}} \succ \mathbf{y} \implies (\tilde{\mathbf{x}}, z_{\hat{p}}) \succsim (\mathbf{y}, z_{\hat{p}}).$$

Construct $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{\hat{p}}$ such that

$$(a) \hat{x}_i = \tilde{x}_i \text{ for all } i \in \{1, \dots, n(\tilde{\mathbf{x}}) - 1\}; \hat{x}_{n(\tilde{\mathbf{x}})} = (\tilde{x}_{n(\tilde{\mathbf{x}})}^w, p_{\tilde{x}}); \hat{x}_{n(\tilde{\mathbf{x}})+1} = (z, \hat{p} + \varepsilon_{\tilde{x}});$$

$$(b) \hat{y}_i = y_i \text{ for all } i \in \{1, \dots, n(\mathbf{y}) - 1\}; \hat{y}_{n(\mathbf{y})} = (y_{n(\mathbf{y})}^w, p_y); \hat{y}_{n(\mathbf{y})+1} = (z, \hat{p} - \varepsilon_y).$$

By construction, $\nu(\hat{\mathbf{x}}) = \nu(\hat{\mathbf{y}}) = \nu(\tilde{\mathbf{x}}, z_{\hat{p}}) = \nu(\mathbf{y}, z_{\hat{p}}) = \hat{\nu}$, $\hat{\mathbf{x}}_{[\cdot]} > (\tilde{\mathbf{x}}, z_{\hat{p}})_{[\cdot]}$ and $\hat{\mathbf{y}}_{[\cdot]} < (\mathbf{y}, z_{\hat{p}})_{[\cdot]}$.

By Axioms 1 and 3,

$$(\tilde{\mathbf{x}}, z_{\hat{p}}) \succsim (\mathbf{y}, z_{\hat{p}}) \implies \hat{\mathbf{x}} \succ \hat{\mathbf{y}}.$$

Also, by construction, $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{\mathbb{Q}_{++}}$.⁶ Hence, by Lemma 2, we know that

$$\hat{\mathbf{x}} \succ \hat{\mathbf{y}} \iff \int_0^{\hat{\nu}'} e^{-\delta\rho} u(\hat{\mathbf{x}}_{[\rho]}) d\rho > \int_0^{\hat{\nu}'} e^{-\delta\rho} u(\hat{\mathbf{y}}_{[\rho]}) d\rho$$

Let \tilde{r} be the rank of $x_{n(\tilde{\mathbf{x}})}$ in $\tilde{\mathbf{x}}$ and r be the rank of $y_{n(\mathbf{y})}$ in \mathbf{y} . By definition of $\hat{\mathbf{x}}_{[\cdot]}$ and $\hat{\mathbf{y}}_{[\cdot]}$, we have:

$$\begin{aligned} & \int_0^{\hat{\nu}} e^{-\delta\rho} u(\hat{\mathbf{x}}_{[\rho]}) d\rho \\ = & \int_0^{\rho_{\tilde{r}} - \varepsilon_{\tilde{x}}} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho + e^{\delta\varepsilon_{\tilde{x}}} \int_{\rho_{\tilde{r}}}^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho_{\tilde{r}}]}) d\rho \\ & + \int_{\nu - \varepsilon_{\tilde{x}}}^{\nu} e^{-\delta\rho} u(z) d\rho + \int_{\nu}^{\hat{\nu}} e^{-\delta\rho} u(z) d\rho \\ = & \int_0^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho + (e^{\delta\varepsilon_{\tilde{x}}} - 1) \int_{\rho_{\tilde{r}}}^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho_{\tilde{r}}]}) d\rho - \int_{\rho_{\tilde{r}} - \varepsilon_{\tilde{x}}}^{\rho_{\tilde{r}}} e^{-\delta\rho} u(x_{n(\tilde{\mathbf{x}})}^w) d\rho \\ & + \int_{\nu - \varepsilon_{\tilde{x}}}^{\nu} e^{-\delta\rho} u(z) d\rho + \int_{\nu}^{\hat{\nu}} e^{-\delta\rho} u(z) d\rho \\ = & \int_0^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho + \int_{\nu}^{\hat{\nu}} e^{-\delta\rho} u(z) d\rho \\ & + (e^{\delta\varepsilon_{\tilde{x}}} - 1) \left(\int_{\rho_{\tilde{r}}}^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho_{\tilde{r}}]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_{\tilde{r}}} u(x_{n(\tilde{\mathbf{x}})}^w)}{\delta} \right) \end{aligned}$$

and likewise

$$\int_0^{\hat{\nu}} e^{-\delta\rho} u(\hat{\mathbf{y}}_{[\rho]}) d\rho$$

⁶Indeed, $\nu - \tilde{x}_{n(\tilde{\mathbf{x}})}^p$ and $\nu - y_{n(\mathbf{y})}^p$ are rational number because all individuals but the last one have rational probabilities of existing.

$$\begin{aligned}
&= \int_0^{\rho_r + \varepsilon_y} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho + e^{-\delta\varepsilon_y} \int_{\rho_r}^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho_r]}) d\rho \\
&\quad - \int_{\nu}^{\nu + \varepsilon_y} e^{-\delta\rho} u(z) d\rho + \int_{\nu}^{\hat{\nu}} e^{-\delta\rho} u(z) d\rho \\
&= \int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho + \int_{\nu}^{\hat{\nu}} e^{-\delta\rho} u(z) d\rho \\
&\quad - (1 - e^{-\delta\varepsilon_y}) \left(\int_{\rho_r}^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho_r]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_r} u(x_n^w(\mathbf{y}))}{\delta} \right)
\end{aligned}$$

To sum up: $\mathbf{x} \succ \mathbf{y} \implies \tilde{\mathbf{x}} \succ \mathbf{y} \implies (\tilde{\mathbf{x}}, (z)_{\hat{p}}) \succsim (\mathbf{y}, (z)_{\hat{p}}) \implies \hat{\mathbf{x}} \succ \hat{\mathbf{y}} \implies$

$$\begin{aligned}
&\int_0^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho + (e^{\delta\varepsilon_{\tilde{x}}} - 1) \left(\int_{\rho_{\tilde{r}}}^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho_{\tilde{r}}]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_{\tilde{r}}} u(\tilde{x}_n^w(\tilde{\mathbf{x}}))}{\delta} \right) > \\
&\int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho - (1 - e^{-\delta\varepsilon_y}) \left(\int_{\rho_r}^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho_r]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_r} u(x_n^w(\mathbf{y}))}{\delta} \right).
\end{aligned}$$

This implication is true for any $(\varepsilon_{\tilde{x}}, \varepsilon_y) \in \mathbb{R}_{++}^2$ as defined above. Since rational number are dense in the real number, it is possible to find a sequence of $((\varepsilon_{\tilde{x}}, \varepsilon_y)) \in (\mathbb{R}^2)^{\mathbb{N}}$ such that each of $\varepsilon_{\tilde{x}}$ and ε_y tends to zero. Hence:

$$\mathbf{x} \succ \mathbf{y} \implies \int_0^{\nu} e^{-\delta\rho} u(\tilde{\mathbf{x}}_{[\rho]}) d\rho \geq \int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho.$$

Figure 1 illustrates the construction of the different allocations involved in Step 1.

Step 2: $\int_0^{\nu} e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho > \int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho \implies \mathbf{x} \succ \mathbf{y}$.

Assume that $\int_0^{\nu} e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho > \int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho$. Since rational number are dense in real numbers, it is possible to find $(\varepsilon_x, \varepsilon_y) \in (0, 1)^2$ such that $p_x = x_{n(\mathbf{x})}^p + \varepsilon_x$ and $p_y = y_{n(\mathbf{y})}^p - \varepsilon_y$ satisfy $(p_x, p_y) \in \mathbb{Q}_{++}^2$, and:

$$\begin{aligned}
&\int_0^{\nu} e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho - (1 - e^{-\delta\varepsilon_x}) \left(\int_{\rho_{\tilde{r}}}^{\nu} e^{-\delta\rho} u(\mathbf{x}_{[\rho_{\tilde{r}}]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_{\tilde{r}}} u(x_n^w(\mathbf{x}))}{\delta} \right) > \\
&\int_0^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho + (e^{\delta\varepsilon_y} - 1) \left(\int_{\rho_r}^{\nu} e^{-\delta\rho} u(\mathbf{y}_{[\rho_r]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_r} u(x_n^w(\mathbf{y}))}{\delta} \right)
\end{aligned}$$

where \tilde{r} is the rank of $x_{n(\mathbf{x})}$ in \mathbf{x} , r is the rank of $y_{n(\mathbf{y})}$ in \mathbf{y} , and $z = \max\{\tilde{\mathbf{x}}_{[\nu]}, \mathbf{y}_{[\nu]}\}$.

Let $\varepsilon_x < \hat{p} < 1$ be such that $\hat{\nu} = \nu + \hat{p}$ satisfies $\hat{\nu} \in \mathbb{Q}_{++}$. We can construct $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{\hat{\nu}}$ in the following way:

- (a) $\hat{x}_i = x_i$ for all $i \in \{1, \dots, n(\mathbf{x}) - 1\}$; $\hat{x}_{n(\mathbf{x})} = (x_{n(\mathbf{x})}^w, p_x)$; $\hat{x}_{n(\mathbf{x})+1} = (z, \hat{p} - \varepsilon_x)$;
- (b) $\hat{y}_i = y_i$ for all $i \in \{1, \dots, n(\mathbf{y}) - 1\}$; $\hat{y}_{n(\mathbf{y})} = (y_{n(\mathbf{y})}^w, p_y)$; $\hat{y}_{n(\mathbf{y})+1} = (z, \hat{p} + \varepsilon_y)$;

so that $\hat{\mathbf{x}}, \hat{\mathbf{y}} \in \mathbf{X}_{\mathbb{Q}_{++}}$, $\nu(\hat{\mathbf{x}}) = \nu(\hat{\mathbf{y}}) = \nu(\mathbf{x}, (z)_{\hat{p}}) = \nu(\mathbf{y}, (z)_{\hat{p}}) = \hat{\nu}$, $\hat{\mathbf{x}}_{[\cdot]} < (\mathbf{x}, (z)_{\hat{p}})_{[\cdot]}$ and $\hat{\mathbf{y}}_{[\cdot]} > (\mathbf{y}, (z)_{\hat{p}})_{[\cdot]}$ (see Figure 2).

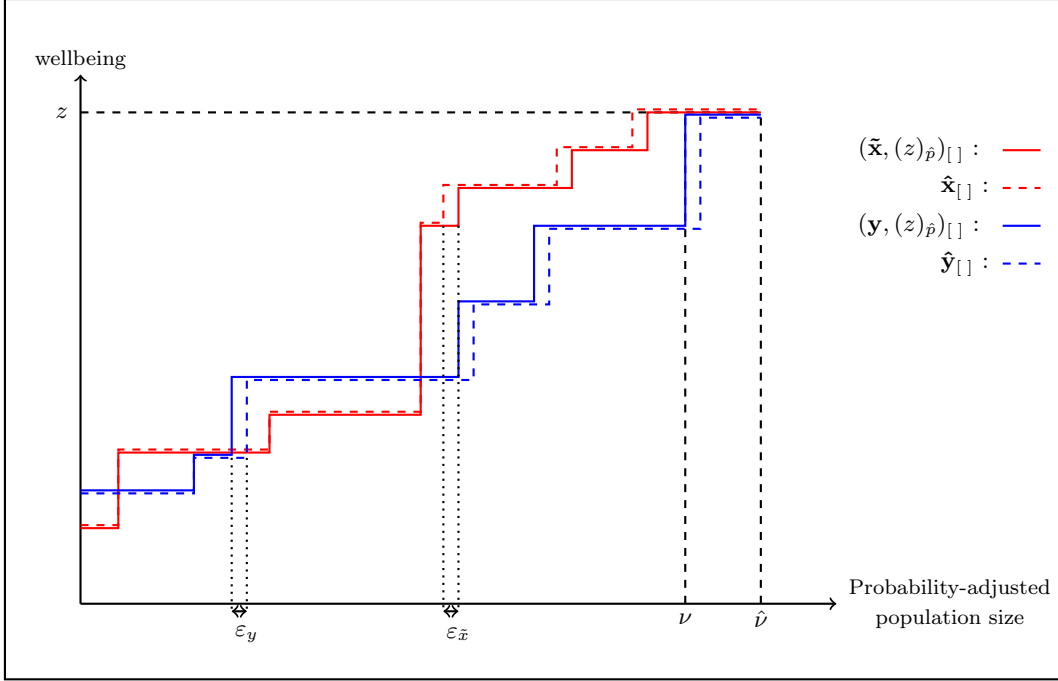


Figure 1: Allocations involved in Step 1 of the proof

By Lemma 2 and by construction of $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$,

$$\begin{aligned}
& \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho - (1 - e^{-\delta\varepsilon_x}) \left(\int_{\rho_r}^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho_r]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_r} u(x_n^w(\mathbf{x}))}{\delta} \right) > \\
& \int_0^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho + (e^{\delta\varepsilon_y} - 1) \left(\int_{\rho_{r'}}^\nu e^{-\delta\rho} u(\mathbf{y}_{[\rho_{r'}]}) d\rho + \frac{e^{-\delta\nu} u(z) - e^{-\delta\rho_{r'}} u(x_n^w(\mathbf{y}))}{\delta} \right) \\
\implies & \int_0^{\hat{\nu}} e^{-\delta\rho} u(\hat{\mathbf{x}}_{[\rho]}) d\rho > \int_0^{\hat{\nu}} e^{-\delta\rho} u(\hat{\mathbf{y}}_{[\rho]}) d\rho \implies \hat{\mathbf{x}} \succ \hat{\mathbf{y}}. \text{ And by Axioms 1, 3 and 4,} \\
& \hat{\mathbf{x}} \succ \hat{\mathbf{y}} \implies (\mathbf{x}, (z)_{\hat{p}}) \succ (\mathbf{y}, (z)_{\hat{p}}) \implies \mathbf{x} \succ \mathbf{y}. \blacksquare
\end{aligned}$$

Finally, we extend the representation to the entire domain \mathbf{X} of all finite allocations (thereby also considering allocations with different probability-adjusted population sizes) by showing that any finite allocation \mathbf{x} can be made as bad as an allocation where all individuals are at the critical level c by adding sufficiently many people at a low wellbeing level z , and thus indifferent to an egalitarian allocation where each individual's wellbeing equals $x \leq c$. This allows us to apply Axiom 6, thereby completing the demonstration of the result that statement (1) of Theorem 1 implies statement (2).

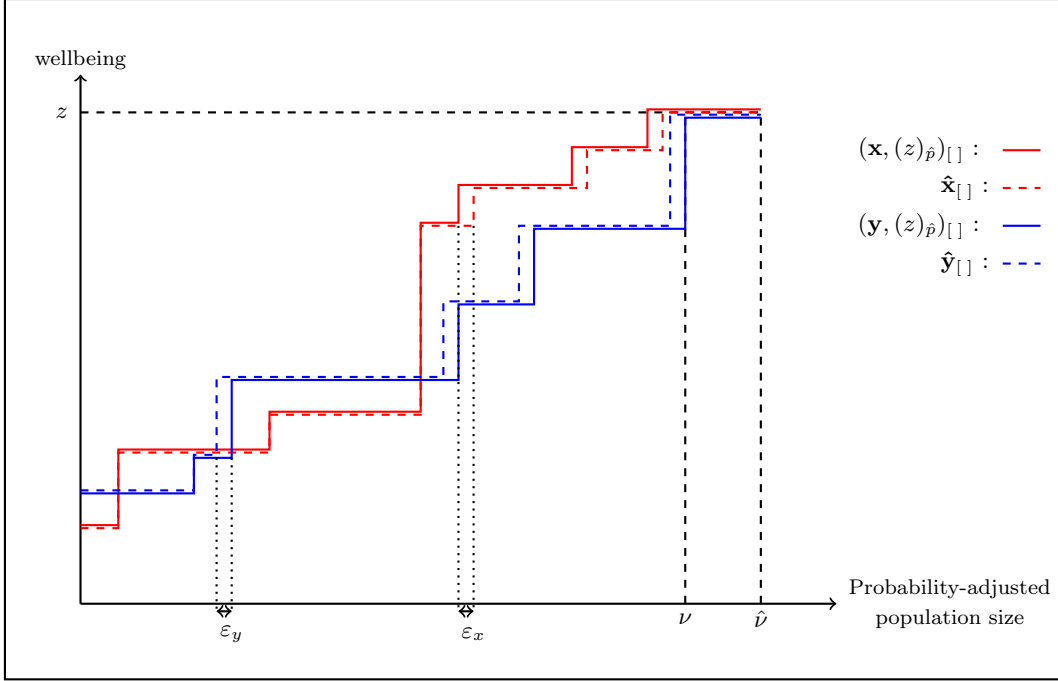


Figure 2: Allocations involved in Step 2 of the proof

Lemma 4 *If Axioms 1-7 hold, then there exists $c \in \mathbb{R}_+$, $\delta \in \mathbb{R}_{++}$ and a continuous and increasing function $u : \mathbb{R} \rightarrow \mathbb{R}$ such that for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$,*

$$\mathbf{x} \succsim \mathbf{y} \iff \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} (u(\mathbf{x}_{[\rho]}) - u(c)) d\rho \geq \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} (u(\mathbf{y}_{[\rho]}) - u(c)) d\rho. \quad (\text{A4})$$

Proof. *Step 1: Representation when well-being does not exceed c .*

Let $c \in \mathbb{R}_+$ be the critical level parameter defined in Axiom 6.

Assume that $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ are such that $\mathbf{x}_{[\nu(\mathbf{x})]} \leq c$ and $\mathbf{y}_{[\nu(\mathbf{y})]} \leq c$. If $\nu(\mathbf{x}) = \nu(\mathbf{y})$, then equivalence (A4) follows from Lemma 3. Therefore, assume that $\nu(\mathbf{x}) < \nu(\mathbf{y})$ (the case $\nu(\mathbf{x}) > \nu(\mathbf{y})$ can be treated similarly). Let $k := \min_{\ell \in \mathbb{N}} \{\ell : (\nu(\mathbf{y}) - \nu(\mathbf{x}))/\ell \leq 1\}$ and $p = (\nu(\mathbf{y}) - \nu(\mathbf{x}))/k$. Then, by k applications of Axiom 6, using Axiom 5 repeatedly to ensure that the allocation is in \mathbf{X}_ν when Axiom 6 is applied, $\mathbf{x} \sim (\mathbf{x}, (c)_{kp})$. By Axiom 1 and Lemma 3:

$$\begin{aligned} \mathbf{x} \succsim \mathbf{y} &\iff (\mathbf{x}, (c)_{kp}) \succsim \mathbf{y} \\ &\iff \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho + \int_{\nu(\mathbf{x})}^{\nu(\mathbf{y})} e^{-\delta\rho} u(c) d\rho \geq \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho \end{aligned}$$

$$\iff \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} (u(\mathbf{x}_{[\rho]}) - u(c)) d\rho \geq \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} (u(\mathbf{y}_{[\rho]}) - u(c)) d\rho.$$

Step 2: Equally distributed equivalent.

For any $\nu \in \mathbb{R}_{++}$ and $\mathbf{x} \in \mathbf{X}_\nu$, let the ν -equally distributed equivalent of \mathbf{x} , denoted $e_\nu(\mathbf{x})$, be $x \in \mathbb{R}$ such that $(x)_\nu \sim \mathbf{x}$. Axioms 1–3 imply that $e_\nu : \mathbf{X}_\nu \rightarrow \mathbb{R}$ is well-defined. By Lemma 3, and since Axioms 1–7 hold, it is defined as follows:

$$e_\nu(\mathbf{x}) = u^{-1} \left(\frac{\delta}{1-e^{-\delta\nu}} \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho \right).$$

Let $\mathbf{x} \in \mathbf{X}_\nu$ and $z < \min\{\mathbf{x}_{[0]}, c\}$, leading to the following expression for $k \in \mathbb{N}$:

$$\begin{aligned} e_{\nu+k}(\mathbf{x}, (z)_k) &= u^{-1} \left(\frac{\delta}{1-e^{-\delta(\nu+k)}} \left(\int_0^k e^{-\delta\rho} u(z) d\rho + \int_k^{\nu+k} e^{-\delta\rho} u(\mathbf{x}_{[\rho-k]}) d\rho \right) \right) \\ &= u^{-1} \left(\frac{1-e^{-\delta k}}{1-e^{-\delta(\nu+k)}} u(z) + \frac{e^{-\delta k} - e^{-\delta(\nu+k)}}{1-e^{-\delta(\nu+k)}} \cdot \frac{\delta}{1-e^{-\delta\nu}} \int_0^\nu e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho \right). \end{aligned}$$

Write $a(k) := (1 - e^{-\delta k}) / (1 - e^{-\delta(\nu+k)})$; note that $a : \mathbb{N} \rightarrow \mathbb{R}$ is an increasing function of k converging to 1. Since $z < \mathbf{x}_{[0]} \leq e_\nu(\mathbf{x})$ and

$$e_{\nu+k}(\mathbf{x}, (z)_k) = u^{-1} (a(k)u(z) + (1 - a(k))u(e_\nu(\mathbf{x}))),$$

it follows that $e_{\nu+k}(\mathbf{x}, (z)_k)$ is a decreasing function of k converging to z as k approaches infinity. As $z < c$, we deduce that, for any $\mathbf{x} \in \mathbf{X}$, there exists $K(\mathbf{x}) \in \mathbb{N}$ such that, for all $k \geq K(\mathbf{x})$, $e_{\nu(\mathbf{x})+k}(\mathbf{x}, (z)_k) \leq c$.

Step 3: Conclusion. For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, choose z such that $z < \min\{\mathbf{x}_{[0]}, \mathbf{y}_{[0]}, z\}$. Let $\ell = \max\{K(\mathbf{x}), K(\mathbf{y})\}$, $x = e_{\nu(\mathbf{x})+\ell}(\mathbf{x}, (z)_\ell)$ and $y = e_{\nu(\mathbf{y})+\ell}(\mathbf{y}, (z)_\ell)$. By definition, $(\mathbf{x}, (z)_\ell) \sim (x)_{\nu(\mathbf{x})+\ell}$, $(\mathbf{y}, (z)_\ell) \sim (y)_{\nu(\mathbf{y})+\ell}$, $x \leq c$ and $y \leq c$. Hence, by repeated applications of Axioms 1 and 5, and by Step 1:

$$\begin{aligned} \mathbf{x} \succsim \mathbf{y} &\iff (\mathbf{x}, (z)_\ell) \succsim (\mathbf{y}, (z)_\ell) \\ &\iff (x)_{\nu(\mathbf{x})+\ell} \succsim (y)_{\nu(\mathbf{y})+\ell} \\ &\iff \int_0^{\nu(\mathbf{x})+\ell} e^{-\delta\rho} (u(x) - u(c)) d\rho \geq \int_0^{\nu(\mathbf{y})+\ell} e^{-\delta\rho} (u(y) - u(c)) d\rho. \end{aligned}$$

However, by the definition of equally distributed equivalents,

$$\begin{aligned} \int_0^{\nu(\mathbf{x})+\ell} e^{-\delta\rho} u(x) d\rho &= \int_0^\ell e^{-\delta\rho} u(z) d\rho + e^{-\delta\ell} \int_0^{\nu(\mathbf{x})} e^{-\delta\rho} u(\mathbf{x}_{[\rho]}) d\rho, \\ \int_0^{\nu(\mathbf{y})+\ell} e^{-\delta\rho} u(y) d\rho &= \int_0^\ell e^{-\delta\rho} u(z) d\rho + e^{-\delta\ell} \int_0^{\nu(\mathbf{y})} e^{-\delta\rho} u(\mathbf{y}_{[\rho]}) d\rho, \end{aligned}$$

Thereby we obtain equivalence (A4). ■