Revisiting variable-value population principles

Walter Bossert
University of Montreal

Susumu Cato
University of Tokyo

Kohei Kamaga
Sophia University

September 27, 2021

E-21-007

Revisiting variable-value population principles[*]

Walter Bossert
Centre Interuniversitaire de Recherche
en Économie Quantitative (CIREQ)
University of Montreal
P.O. Box 6128, Station Downtown
Montreal QC H3C 3J7
Canada
walter.bossert@videotron.ca

Susumu Cato
Institute of Social Science
University of Tokyo
7-3-1, Hongo, Bunkyo-ku
Tokyo 113-0033
Japan
susumu.cato@gmail.com

Kohei Kamaga
Faculty of Economics
Sophia University
7-1, Kioi-cho, Chiyoda-ku
Tokyo 102-8554
Japan
kohei.kamaga@sophia.ac.jp

This version: September 21, 2021

Abstract. This paper examines a general class of variable-value population principles. Our particular focus is on the extent to which such principles can avoid the repugnant and sadistic conclusions, two well-established desiderata in population ethics. We show that if a relatively mild limit property is imposed, avoidance of the repugnant conclusion implies the sadistic conclusion. This result generalizes earlier observations by showing that they apply to a substantially larger class of principles. Our second theorem states that, under the limit property, the axiom of mere addition (a condition that is closely related to avoidance of the sadistic conclusion) also conflicts with avoidance of the repugnant conclusion. The latter result is a consequence of a similar observation that appears in the earlier literature. *Journal of Economic Literature* Classification Nos.: D31, D63.

Keywords: variable-value principles, repugnant conclusion, sadistic conclusion, mere addition, number-dampened utilitarianism, critical-level utilitarianism, rank-discounted utilitarianism.

1 Introduction

Ever since its conception, Parfit's (1976, 1982, 1984) repugnant conclusion has remained a fundamental criterion used to assess the ethical merits of population principles. Roughly speaking, the repugnant conclusion states that population size can always be used as a substitute for quality of life, even if the lives of those alive are barely worth living. Initially intended as an argument against total utilitarianism, avoidance of the repugnant conclusion has become a condition that, although not universally accepted, is widely endorsed by a substantial number of scholars in the field. Authors who suggest that the repugnant conclusion may be acceptable include Mackie (1985), Hare (1988), Tännsjö (2002), Huemer (2008), Spears and Budolfson (2021), and Gustafsson (forthcoming). Attempts to avoid the repugnant conclusion are reviewed by Arrhenius, Ryberg, and Tännsjö (2017).

A population principle is expressed by means of an ordering that ranks utility distributions, where utility is a measure of lifetime well-being and utility distributions may be associated with different populations and different population sizes. A life is a neutral life if it is, from the viewpoint of the individual leading it, as good as a life without any experiences. As is common in population ethics, we assume that individual utilities are normalized so that a utility level of zero corresponds to a neutral life. Thus, a life is worth living for a person if it is associated with a positive level of lifetime utility; if lifetime utility is negative, it is not.

The repugnant conclusion is implied if, for any population size n, any arbitrarily high positive utility level ξ, and any arbitrarily low positive utility level ε, there exists a larger population size $m > n$ such that a distribution in which the m people alive experience the (arbitrarily close to zero) utility level ε is better than a distribution in which the n individuals alive are at the (arbitrarily high) utility level ξ. In other words, a principle that implies the repugnant conclusion permits mass poverty to substitute for quality of life—an observation that seems very repugnant indeed.

An early response to Parfit's repugnant conclusion is provided in a fundamental contribution by Hurka (1983). He proposes the notion of a variable-value principle, the purpose of which is to capture the behavior of a total principle (such as total utilitarianism) for small population sizes and that of an average rule for large numbers. Thus, the value assigned to an additional life varies depending on the size of the existing population. The variable-value principles are obtained by multiplying the utility of an equal distribution by a non-decreasing dampening function of population size. As Hurka (1983) demonstrates, if the function is bounded, the repugnant conclusion is indeed avoided: the effect of the total-utilitarian criterion is dampened for large populations by applying this function, and this dampening procedure allows for an escape from the repugnant conclusion. A fully specified variant of these principles appears in Ng (1986) who proposes the number-dampened utilitarian principles. They are defined for all distributions and use the criterion that is obtained by multiplying average utility by the value of a non-decreasing dampening function. The resulting population principle avoids the repugnant conclusion if and only if the dampening function is bounded; see, for example, Blackorby, Bossert, and Donaldson (2005, p. 172).

The class of critical-level utilitarian principles introduced and analyzed by Blackorby

and Donaldson (1984) represents another method of avoiding the repugnant conclusion. A critical level of lifetime well-being is a level of utility such that if a person with that utility level is added to a utility-unaffected population, the expanded distribution is as good as the original. A critical level need not exist for some principles and for some distributions. If a critical level exists, it may depend on the distribution under consideration, and it may not be unique. Critical-level utilitarianism ranks distributions by means of the sum of the differences between individual utilities and a fixed critical level. A special case is total utilitarianism, which is obtained if the critical level is equal to zero—the level of neutrality. The repugnant conclusion is avoided if and only if the critical level is positive; see Blackorby and Donaldson (1984).

An alternative assessment criterion for population principles is introduced by Arrhenius (2000). A principle implies the sadistic conclusion if there is an initial utility distribution such that adding only people with negative levels of well-being (below neutrality) to that distribution is better than adding only individuals with positive utilities (above neutrality).

It is well-known that there is a tension between the two objectives of avoiding the repugnant conclusion and avoiding the sadistic conclusion; this conflict has been analyzed quite thoroughly in the earlier literature for some specific classes of population principles. In particular, Blackorby, Bossert, and Donaldson (2005, p. 164, Theorem 5.4) show that there is no principle with utilitarian fixed-number restrictions that can avoid both conclusions under a mild regularity condition. Other incompatibilities in population ethics are examined by Ng (1989) and by Carlson (1998), for instance.

Arrhenius (2000) shows that a subclass of the number-dampened utilitarian principles leads to the sadistic conclusion. Although he only proves this for a specific functional form of the dampening function, his observation generalizes to all members of the class that employ a bounded function. Thus, in the context of the number-dampened utilitarian principles, the incompatibility between avoidance of the repugnant conclusion and avoidance of the sadistic conclusion is well-established and well-understood. Analogously, no critical-level utilitarian principles can avoid the repugnant conclusion and the sadistic conclusion, as demonstrated by Blackorby, Bossert, and Donaldson (2004); see also Bossert (forthcoming). They show that a critical-level utilitarian principle avoids the sadistic conclusion if and only if the critical level is equal to zero. Combined with the above-mentioned observation that avoidance of the repugnant conclusion is satisfied if and only if the critical level is positive, the impossibility result follows immediately.

In this paper, we extend incompatibility results of this nature by demonstrating that they apply to a considerably more general class of population principles. These principles are what we refer to as the gains-based variable-value principles, and they generalize Hurka's (1983) class of variable-value principles by allowing for a reference level of utility that need not be equal to the level of neutrality. In particular, the principles assess equal distributions by comparing the difference between the equal utility level and a reference level, modified by the value of a non-decreasing dampening function that depends on the population size. Hurka's class of variable-value principles emerges as a special case if the reference level is equal to zero. In addition to the variable-value principles and the critical-level utilitarian principles, numerous others are included in our class. As a prominent example, the rank-discounted reference-level utilitarian principles of Asheim and Zuber (2014) are among

them.

The generality of our observations comes at a price. Because the (gains-based) variable-value principles are defined solely on the basis of equal utility distributions, an additional condition is required; without such a property, avoidance of the sadistic conclusion does not have enough bite if nothing can be said about unequal distributions. This is a feature that our approach shares with some earlier contributions to the literature. The limit property, as we label it, is a weakening of the aggregation property that is employed by Spears and Budolfson (2021). It slightly strengthens the conjunction of the positive limit property and the negative limit property used by Blackorby, Bossert, Donaldson, and Fleurbaey (1998) who consider orderings that are not necessarily anonymous. Our limit property requires that, as the total population becomes large, inequities that may be present in a given distribution eventually become negligible. This means that the condition ensures that equitable assessments are approached in the limit.

We note that all of our observations remain valid if any continuous and increasing transformation that preserves the neutrality normalization is applied to all utility values. The reason why we do not work with the resulting generalized principles is ease of exposition.

Section 2 introduces our class of gains-based variable-value principles, along with some examples. In Section 3, we provide formal definitions of the repugnant and sadistic conclusions and our limit property. Moreover, we state and prove our main result. Section 4 is devoted to a discussion and a related impossibility result obtained for Parfit's (1984) axiom of mere addition in place of the sadistic conclusion. Section 5 concludes.

2 Gains-based variable-value principles

We use individual utility as an indicator of lifetime well-being. Following the standard convention, utilities are normalized in a way such that a value of zero represents a neutral life. A life is neutral if it is, from the viewpoint of the person leading it, neither better nor worse than a life without any experiences.

A utility distribution is a list of individual lifetime utility levels, one level for each person in existence in the requisite distribution. Thus, for a population of $n \in \mathbb{N}$ individuals, a distribution is given by a vector $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$. The vector consisting of $n \in \mathbb{N}$ ones is denoted by $\mathbf{1}_n$.

A population principle is an ordering (that is, a complete and transitive binary relation) R defined on the set of possible distributions $\Omega = \cup_{n \in \mathbb{N}} \mathbb{R}^n$. The statement uRv is interpreted to mean that distribution $u \in \Omega$ is at least as good as distribution $v \in \Omega$ according to the population principle identified by R. We use P and I to denote the asymmetric and symmetric parts of R, that is, uPv indicates that u is better than v, and uIv means that u and v are equally good.

Consider a population principle R and a utility distribution $u \in \Omega$. A utility level $c \in \mathbb{R}$ is a critical level for the principle R and the distribution u if the distribution (u, c) that results from adding a person with lifetime utility c to the distribution u is as good as u. Note that, without further assumptions, a critical level need not exist for all distributions, and it need not be unique if it exists. Moreover, a critical level may depend on the utility

distribution under consideration.

The class of population principles that we focus on is given by the gains-based variable-value principles defined as follows. An ordering R on Ω is a gains-based variable-value principle if there exist a number $\gamma \in \mathbb{R}$ and a non-decreasing function $f \colon \mathbb{N} \to \mathbb{R}_{++}$ such that, for all population sizes $n, m \in \mathbb{N}$ and for all utility levels $a, b \in \mathbb{R}$,

$$a\mathbf{1}_n R b\mathbf{1}_m \;\Leftrightarrow\; f(n)(a - \gamma) \geq f(m)(b - \gamma).$$

The parameter γ is a reference level of utility and f is a dampening function that determines the value attached to an additional person as a function of the population size. The term variable value is motivated by the observation that this value may depend on the number of people already in existence. Thus, gains-based variable-value principles assess equal distributions by comparing the difference between the equal utility level and this reference level, modified by the value of the function f that depends on the population size. This class generalizes the variable-value principles employed by Hurka (1983); the parameter γ is equal to zero in Hurka's account.

As mentioned in the introduction, all of our observations extend easily to the generalized variants of these principles. These generalizations are obtained by applying a continuous and increasing transformation that preserves the zero normalization for neutrality to all utilities. All that needs to be done is to replace utilities by transformed utilities in our arguments.

Gains-based variable-value population principles encompass a plethora of orderings because their definition merely imposes a restriction on comparisons that involve equal utility distributions. Moreover, they generalize the traditional number-dampened principles because they allow for the reference utility γ to differ from the neutrality level zero. As a consequence, they are capable of accommodating critical-level principles such as those introduced by Blackorby and Donaldson (1984) and discussed by Blackorby, Bossert, and Donaldson (2005), among others.

The first of the special cases we mention here is number-dampened utilitarianism; see Hurka (1983) and Ng (1986). An ordering R on Ω is a number-dampened utilitarian principle if there exists a non-decreasing function $f \colon \mathbb{N} \to \mathbb{R}_{++}$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \;\Leftrightarrow\; \frac{f(n)}{n} \sum_{i=1}^{n} u_i \geq \frac{f(m)}{m} \sum_{i=1}^{m} v_i.$$

The reference level of utility γ is equal to zero in this case. If f is equal to the identity mapping (or any increasing linear function), total utilitarianism is obtained. For a constant function f, average utilitarianism results.

A second example is the class of Sider's (1991) principles, also referred to as geometrism. To define these principles, we require more notation. For $n \in \mathbb{N}$ and $u \in \mathbb{R}^n$, let $u_{[\,]}$ denote a non-decreasing rearrangement of the components of u, that is,

$$u_{[1]} \leq \ldots \leq u_{[n]}.$$

4

Analogously, let $u_{()}$ be a non-increasing rearrangement of u so that

$$u_{(1)} \geq \ldots \geq u_{(n)}.$$

For a utility distribution $u \in \Omega$, the number of non-negative components in u is denoted by $n^+(u)$, and the number of negative components is $n^{--}(u)$. Let $u^{(+)}$ be a non-increasing rearrangement of the non-negative components of u so that

$$u^{(+)} = (u^{(+)}_{(1)}, u^{(+)}_{(2)}, \ldots, u^{(+)}_{(n^+(u))}),$$

and, analogously, $u^{[--]}$ is a non-decreasing rearrangement of negative components of u, that is,

$$u^{[--]} = (u^{[--]}_{[1]}, u^{[--]}_{[2]}, \ldots, u^{[--]}_{[n^{--}(u)]}).$$

An ordering R on Ω is geometrist if there exists $\beta \in (0,1)$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \iff \sum_{i=1}^{n^+(u)} \beta^{i-1} u_{(i)} + \sum_{i=1}^{n^{--}(u)} \beta^{i-1} u_{[i]} \geq \sum_{i=1}^{n^+(v)} \beta^{i-1} v_{(i)} + \sum_{i=1}^{n^{--}(v)} \beta^{i-1} v_{[i]}.$$

Sider's geometrist principles can be considered variants of rank-discounted utilitarianism. An ordering R on Ω is rank-discounted utilitarian if there exists $\beta \in (0,1)$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \iff \sum_{i=1}^{n} \beta^{i-1} u_{[i]} \geq \sum_{i=1}^{m} \beta^{i-1} v_{[i]}.$$

As is the case for number-dampened utilitarianism, the reference level of utility for geometrism and for rank-discounted utilitarianism is equal to zero—the level that represents neutrality. If γ is allowed to assume values other than zero, the rank-discounted reference-level utilitarian principles of Asheim and Zuber (2014) result. An ordering R on Ω is rank-discounted reference-level utilitarian if there exist $\beta \in (0,1)$ and $\gamma \in \mathbb{R}$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \iff \sum_{i=1}^{n} \beta^{i}[u_{[i]} - \gamma] \geq \sum_{i=1}^{m} \beta^{i}[v_{[i]} - \gamma].$$

Asheim and Zuber (2014) refer to these principles as rank-discounted critical-level utilitarian principles. We prefer to use the term reference level rather than critical level because the parameter γ is not a critical level in the sense employed here. This is the case because adding a person at the utility level γ to a utility-unaffected population does not necessarily lead to a distribution that is as good as the original. See also Asheim and Zuber (forthcoming) and Pivato (2020) for alternative generalized classes.

The class of critical-level utilitarian principles introduced by Blackorby and Donaldson (1984) and further discussed by Blackorby, Bossert, and Donaldson (2005) represents yet

another special case. An ordering R on Ω is critical-level utilitarian if there exists $\alpha \in \mathbb{R}$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \iff \sum_{i=1}^{n}[u_i - \alpha] \geq \sum_{i=1}^{m}[v_i - \alpha].$$

In this case, the parameter γ in the definition of the gains-based variable-value principles is equal to the constant critical level α, and the function f is any increasing linear function. In contrast to the rank-discounted principles, the parameter $\gamma = \alpha$ is indeed a critical level.

3   The repugnant and sadistic conclusions

This section presents our main result, along with a discussion of some of its features. We show that, under a mild additional restriction, the gains-based variable-value principles that avoid the repugnant conclusion must entail the sadistic conclusion. The additional restriction is needed as a consequence of the general nature of the gains-based variable-value principles. Because these principles are defined by exclusively referring to equal distributions, more structure is needed in order to obtain observations that allow some interpersonal variation in individual utility values.

We begin with a formal statement of Parfit's (1976, 1982, 1984) repugnant conclusion. See, for instance, Spears and Budolfson (2021) for alternative definitions. A population principle R implies the repugnant conclusion if, for all $n \in \mathbb{N}$, for all $\xi \in \mathbb{R}_{++}$, and for all $\varepsilon \in (0, \xi)$, there exists $m \in \mathbb{N}$ with $m > n$ such that $\varepsilon\mathbf{1}_m P \xi\mathbf{1}_n$. The axiom avoidance of the repugnant conclusion is obtained by negating this statement.

Avoidance of the repugnant conclusion. There exist $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, and $\varepsilon \in (0, \xi)$ such that, for all $m \in \mathbb{N}$ with $m > n$,

$$\xi\mathbf{1}_n R \varepsilon\mathbf{1}_m.$$

A population principle implies the sadistic conclusion (Arrhenius, 2000) if adding people with negative utility levels can be better than adding people with positive utility levels to a given population. To provide a formal definition, we use the notation $\Omega_{++} = \cup_{n \in \mathbb{N}}\mathbb{R}^n_{++}$ and $\Omega_{--} = \cup_{n \in \mathbb{N}}\mathbb{R}^n_{--}$ to denote the set of all positive distributions and the set of all negative distributions. Thus, the sadistic conclusion is implied by a population principle R if there exist $u \in \Omega$, $v \in \Omega_{++}$, and $w \in \Omega_{--}$ such that $(u, w)P(u, v)$. Negating this statement, we obtain the following axiom.

Avoidance of the sadistic conclusion. For all $u \in \Omega$, for all $v \in \Omega_{++}$, and for all $w \in \Omega_{--}$, $(u, v)R(u, w)$.

This axiom is called the non-sadism condition by Arrhenius (forthcoming).

As alluded to earlier, we require an additional property to be able to deal with some distributions that do not involve equal utilities for everyone. We employ a variant of the aggregation axiom of Spears and Budolfson (2021).

6

Limit property. For all $n \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, for all $\kappa \in \mathbb{R}$, and for all $\delta \in \mathbb{R}_{++}$, there exists $m^* \in \mathbb{N}$ such that, for all $m \in \mathbb{N}$ with $m > m^*$,

$$(\kappa + \delta)\mathbf{1}_{n+m} \ R \ (\kappa\mathbf{1}_m, u) \ R \ (\kappa - \delta)\mathbf{1}_{n+m} \ \ \text{and} \ \ (\kappa + \delta)\mathbf{1}_{n+m} \ R \ (u, \kappa\mathbf{1}_m) \ R \ (\kappa - \delta)\mathbf{1}_{n+m}.$$

Spears and Budolfson's (2021) axiom of aggregation is obtained by replacing the at-least-as-good relation with betterness in the consequent of the limit property regarding the utility distribution $(\kappa\mathbf{1}_m, u)$. Thus, in the presence of the anonymity axiom that requires any distribution to be declared as good as any of its permutations, Spears and Budolfson's aggregation axiom implies the limit property; indeed, anonymity is one of their basic axioms. Moreover, we note that Spears and Budolfson's aggregation axiom is a strengthening of the positive limit property and its negative counterpart introduced by Blackorby, Bossert, Donaldson, and Fleurbaey (1998). The limit property is also a strengthening of their positive and negative limit properties. The limit property is appealing because it requires that, as the total population becomes large, any inequities that may be present in an original distribution $u \in \mathbb{R}^n$ are trumped by the inequalities $\kappa + \delta > \kappa > \kappa - \delta$. Thus, the condition pushes towards equitable assessments in the limit.

Our main result establishes the fundamental incompatibility of avoidance of the repugnant conclusion and avoidance of the sadistic conclusion for the class of gains-based variable-value principles.

Theorem 1. *Let R be a gains-based variable-value principle that satisfies the limit property. If R avoids the repugnant conclusion, then R implies the sadistic conclusion.*

We use four auxiliary results to prove this theorem. The first of these generalizes an analogous observation of Blackorby and Donaldson (1984) for the critical-level utilitarian principles to our class. It does not rely on the limit property.

Lemma 1. *Let R be a gains-based variable-value principle. If $\gamma > 0$, then R avoids the repugnant conclusion.*

Proof. Assume that $\gamma > 0$. Let $\xi = \gamma$ and $\varepsilon \in (0, \xi) = (0, \gamma)$. Furthermore, let $n \in \mathbb{N}$ be arbitrary. Because f is positive-valued, it follows that

$$f(n)(\xi - \gamma) = 0 > f(m)(\varepsilon - \gamma)$$

for all $m \in \mathbb{N}$ with $m > n$ so that, by definition of the gains-based variable-value principles, $\xi\mathbf{1}_n P \varepsilon\mathbf{1}_m$ and hence $\xi\mathbf{1}_n R \varepsilon\mathbf{1}_m$ for all $m \in \mathbb{N}$ with $m > n$. ∎

The following lemma also has its foundation in an earlier result. As stated in Blackorby, Bossert, and Donaldson (2005, p. 172), a number-dampened utilitarian ordering avoids the repugnant conclusion if and only if the function f is bounded. This observation readily generalizes to our class of principles. Again, the limit property is not required in this result.

Lemma 2. *Let R be a gains-based variable-value principle. If $\gamma \leq 0$, then R avoids the repugnant conclusion if and only if f is bounded.*

Proof. 'If.' Assume that $\gamma \leq 0$ and f is bounded. To prove that R avoids the repugnant conclusion, we have to establish the existence of $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, and $\varepsilon \in (0, \xi)$ such that, for all $m \in \mathbb{N}$ with $m > n$,

$$\xi \mathbf{1}_n R \varepsilon \mathbf{1}_m.$$

Let $\xi \in \mathbb{R}_{++}$ and $\varepsilon \in (0, \xi)$. Since $\gamma \leq 0$, it follows that the differences $\xi - \gamma$ and $\varepsilon - \gamma$ are positive and, moreover, that

$$\frac{\xi - \gamma}{\varepsilon - \gamma} > 1.$$

Because f is non-decreasing and bounded, $f(n)$ converges to a finite limit f^* as n approaches infinity. Thus, there exists $n \in \mathbb{N}$ such that

$$\frac{f^*}{f(n)} < \frac{\xi - \gamma}{\varepsilon - \gamma}.$$

Again using the non-decreasingness of f, it follows that $f(n) \leq f(m) \leq f^*$ and hence

$$\frac{f(m)}{f(n)} \leq \frac{f^*}{f(n)} < \frac{\xi - \gamma}{\varepsilon - \gamma}$$

for all $m \in \mathbb{N}$ with $m > n$. Because f is positive-valued, this implies that

$$f(n)(\xi - \gamma) > f(m)(\varepsilon - \gamma)$$

for all $m \in \mathbb{N}$ such that $m > n$. By definition of the gains-based variable-value principles, it follows that $\xi \mathbf{1}_n P \varepsilon \mathbf{1}_m$ and hence $\xi \mathbf{1}_n R \varepsilon \mathbf{1}_m$, as was to be shown.

'Only if.' Assume that f is unbounded. Let $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, and $\varepsilon \in (0, \xi)$. Since $\gamma \leq 0$, it follows that the differences $\varepsilon - \gamma$ and $\xi - \gamma$ are positive. Because f is unbounded and non-decreasing, there exists a population size $m \in \mathbb{N}$ with $m > n$ such that

$$f(m)(\varepsilon - \gamma) > f(n)(\xi - \gamma).$$

By definition of the gains-based variable-value principles, this inequality is equivalent to the statement $\varepsilon \mathbf{1}_m P \xi \mathbf{1}_n$ so that the repugnant conclusion is implied. ∎

The remaining two lemmas focus on the sadistic conclusion. Both of them employ the limit property. First, we consider the case of a positive parameter value γ.

Lemma 3. *Let R be a gains-based variable-value principle that satisfies the limit property. If $\gamma > 0$, then R implies the sadistic conclusion.*

Proof. Assume that $\gamma > 0$. Let $a, b \in \mathbb{R}_{++}$ and $c \in \mathbb{R}_{--}$ be such that $a > \gamma > b > 0 > c$. We first show that there exists $m^* \in \mathbb{N}$ such that, for all $m \in \mathbb{N}$ with $m > m^*$,

$$(a\mathbf{1}_m, c)P(\gamma, \gamma). \tag{1}$$

Let $\delta \in \mathbb{R}_{++}$ be such that $a - \delta > \gamma$. By the limit property, there exists $m^* \in \mathbb{N}$ such that, for all $m \in \mathbb{N}$ with $m > m^*$,

$$(a\mathbf{1}_m, c)R(a - \delta)\mathbf{1}_{m+1}.$$

For all $m > m^*$, we obtain

$$f(m + 1)[(a - \delta) - \gamma] > 0 = f(2)(\gamma - \gamma).$$

Thus, it follows that

$$(a - \delta)\mathbf{1}_{m+1}P(\gamma, \gamma).$$

By the transitivity of R, we obtain (1).

Now assume that $m \in \mathbb{N}$ is such that (1) is satisfied. Let $\delta \in \mathbb{R}_{++}$ be such that $b + \delta < \gamma$. The limit property implies that there exists $n^* \in \mathbb{N}$ such that, for all $n \in \mathbb{N}$ with $n > n^*$,

$$(b + \delta)\mathbf{1}_{n+m}R(a\mathbf{1}_m, b\mathbf{1}_n).$$

Note that, for any $n > n^*$, we obtain

$$(\gamma, \gamma)P(b + \delta)\mathbf{1}_{n+m}$$

because $b + \delta - \gamma < 0$ and

$$f(2)(\gamma - \gamma) = 0 > f(n + m)(b + \delta - \gamma).$$

By the transitivity of R, we obtain

$$(\gamma, \gamma)P(a\mathbf{1}_m, b\mathbf{1}_n).$$

From (1) and transitivity, it follows that

$$(a\mathbf{1}_m, c)P(a\mathbf{1}_m, b\mathbf{1}_n).$$

This means that R implies the sadistic conclusion. ∎

Lemma 4. *Let R be a gains-based variable-value principle that satisfies the limit property. If $\gamma \leq 0$ and f is bounded, then R implies the sadistic conclusion.*

Proof. Assume that $\gamma \leq 0$ and f is bounded. Let $a, b \in \mathbb{R}_{++}$ and $c \in \mathbb{R}_{--}$ be such that $a > b > 0 > c$. Since f is non-decreasing and bounded, it converges to a finite limit f^* as n approaches infinity. Define $\overline{f} = f^*/f(1)$, and let $\delta \in \mathbb{R}_{++}$ be such that

$$a > a - \delta > b.$$

The limit property implies that there exists $\overline{m} \in \mathbb{N}$ such that, for all $m > \overline{m}$,

$$(a + \delta)\mathbf{1}_{m+1}R(a\mathbf{1}_m, c)R(a - \delta)\mathbf{1}_{m+1}.$$

9

Since $a - \delta - \gamma > b - \gamma > 0$, the definition of $\overline{f}$ implies that there exists $\overline{n} > 0$ such that, for all $n > \overline{n}$,

$$\frac{f(n)}{f(1)}[(a - \delta) - \gamma] > \overline{f}[b - \gamma] > 0.$$

Next, observe that there exists a sufficiently small $\varepsilon \in \mathbb{R}_{++}$ such that, for all $n > \overline{n}$,

$$\frac{f(n)}{f(1)}[(a - \delta) - \gamma] > \overline{f}[(b + \varepsilon) - \gamma] > 0.$$

From the limit property, there exists $n^* > \overline{n}$ such that, for all $n > n^*$,

$$(b + \varepsilon)\mathbf{1}_{\overline{m}+1+n} R(a\mathbf{1}_{\overline{m}+1}, b\mathbf{1}_n) R(b - \varepsilon)\mathbf{1}_{\overline{m}+1+n}.$$

Let $m^* > \max\{n^*, \overline{m}\}$. Then, from the above argument, it follows that

$$(a\mathbf{1}_{m^*}, c)R(a - \delta)\mathbf{1}_{m^*+1}$$

and

$$(b + \varepsilon)\mathbf{1}_{m^*+n^*} R(a\mathbf{1}_{m^*}, b\mathbf{1}_{n^*}).$$

Furthermore, since

$$0 < f(1) \le f(n^*) \le f(m^* + 1) \le f(m^* + n^*) \le f^* = f(1)\overline{f},$$

it follows that

$$
\begin{aligned}
& f(m^* + 1)[(a - \delta) - \gamma] - f(m^* + n^*)[(b + \varepsilon) - \gamma] \\
& \ge f(n^*)[(a - \delta) - \gamma] - f(1)\overline{f}[(b + \varepsilon) - \gamma] \\
& = f(1)\left[\frac{f(n^*)}{f(1)}[(a - \delta) - \gamma] - \overline{f}[(b + \varepsilon) - \gamma] \right] \\
& > 0.
\end{aligned}
$$

Thus, we obtain

$$(a - \delta)\mathbf{1}_{m^*+1} P(b + \varepsilon)\mathbf{1}_{m^*+n^*}.$$

By the transitivity of R, it follows that

$$(a\mathbf{1}_{m^*}, c)P(a\mathbf{1}_{m^*}, b\mathbf{1}_{n^*})$$

which means that R implies the sadistic conclusion. ■

Proof of Theorem 1. By Lemmas 1 and 2, R avoids the repugnant conclusion if and only if $\gamma > 0$ or $\gamma \le 0$ and f is bounded. If $\gamma > 0$, Lemma 3 implies that R entails the sadistic conclusion. If $\gamma \le 0$ and f is bounded, Lemma 4 implies that R entails the sadistic conclusion. ■

In Theorem 1, we restrict attention to gains-based variable-value principles. This raises the question whether the impossibility result is valid for a more general class of population principles. Consider the principle that is defined by letting

$$uPv \ \text{ if } u \in \mathbb{R} \text{ and } v \in \mathbb{R}^m \text{ with } m \in \mathbb{N} \setminus \{1\},$$

and

$$uRv \ \Leftrightarrow \ \sum_{i=1}^n u_i \geq \sum_{i=1}^m v_i$$

for all remaining $n, m \in \mathbb{N}$, $u \in \mathbb{R}^n$, and $v \in \mathbb{R}^m$. According to this principle, all one-person distributions are better than all distributions with more than on individual alive, and all other comparisons are performed by means of the total-utilitarian criterion. This principle satisfies avoidance of the repugnant conclusion, avoidance of the sadistic conclusion, and the limit property. Moreover, it satisfies the fixed-population-size properties of anonymity and strong Pareto; the latter requires that uPv for all $n \in \mathbb{N}$ and for all $u, v \in \mathbb{R}^n$ such that $u_i \geq v_i$ for all $i \in \{1, \ldots, n\}$ with at least one strict inequality. It is clear that this example is not a gains-based variable-value principle because there cannot exist a dampening function f that is non-decreasing; note that the value of f at population size one would have to exceed the values of f at all other population sizes, which is not possible if the monotonicity property of f is to be respected. Thus, the result of Theorem 1 is rather tight.

4 Discussion

As noted earlier, numerous population principles are members of the class of gains-based variable-value principles that satisfy the limit property, among them number-dampened utilitarianism, critical-level utilitarianism, and rank-discounted reference-level utilitarianism. Therefore, these population principles necessarily face the tension between the repugnant and sadistic conclusions. With regard to the case of number-dampened utilitarianism, our result shows that Arrhenius's (2000) proof using the geometric case of f is valid generally without specifying the function f.

Asheim and Zuber (2014) show that rank-discounted reference-level utilitarianism can avoid the repugnant conclusion and what they label the very sadistic conclusion. According to the very sadistic conclusion, there exist a positive utility distribution u and a negative utility distribution v such that v is better than u. Our result implies that although rank-discounted reference-level utilitarianism can satisfy both avoidance of the repugnant conclusion and avoidance of the very sadistic conclusion, it cannot satisfy both avoidance of the repugnant conclusion and avoidance of the sadistic conclusion.

Clearly, there are gains-based variable-value population principles that do not satisfy the limit property. For example, Sider's (1991) principles do not satisfy the limit property. To illustrate this observation, consider the distributions $(b\mathbf{1}_n, a)$ and $(b + \delta)\mathbf{1}_{n+1}$, where $a, b \in \mathbb{R}_{++}$ are such that $a > b$ and $\delta \in \mathbb{R}_{++}$. According to the definition of Sider's principles, it follows that

$$(b+\delta)\mathbf{1}_{n+1}P(b\mathbf{1}_n, a) \ \Leftrightarrow \ (b+\delta)\frac{(1-\beta^{n+1})}{1-\beta} > a + b\beta\frac{1-\beta^n}{1-\beta}.$$

The limit property requires that, for all $\delta \in \mathbb{R}_{++}$, there exists $n^* \in \mathbb{N}$ such that

$$(b + \delta)\frac{(1 - \beta^{n+1})}{1 - \beta} > a + b\beta\frac{1 - \beta^n}{1 - \beta}$$

for all $n \in \mathbb{N}$ such that $n > n^*$. This can only be true if

$$(b + \delta) > (1 - \beta)b + \beta a.$$

However, this inequality is not valid for sufficiently small values of δ. Therefore, the limit property is not satisfied.

Another class of gains-based variable-value principles that do not satisfy the limit property consists of what we label the variable-value maximin principles. The ordering R is a variable-value maximin principle if there exist a non-decreasing and bounded function $f: \mathbb{N} \to \mathbb{R}_{++}$ such that, for all $n, m \in \mathbb{N}$, for all $u \in \mathbb{R}^n$, and for all $v \in \mathbb{R}^m$,

$$uRv \iff f(n)u_{(n)} \geq f(m)v_{(m)}.$$

Alternative maximin-based principles are examined by Bossert (1990), Blackorby, Bossert, and Donaldson (2005), and Zuber (2018). To see that the variable-value maximin principles do not satisfy the limit property, consider the distributions $(b\mathbf{1}_n, c)$ and $(b - \delta)\mathbf{1}_{n+1}$, where $b, c, \delta \in \mathbb{R}_{++}$ are such that $b > c$ and $\delta < b - c$. Using the definition of variable-value maximin, it follows that

$$(b - \delta)\mathbf{1}_{n+1}P(b\mathbf{1}_n, c)$$

for all $n \in \mathbb{N}$ because $b - \delta > c$. This is incompatible with the limit property.

Sider's principles and the variable-value maximin principles avoid both the repugnant and sadistic conclusions. This observation establishes that an axiom such as the limit property is indeed required for our theorem.

Parfit's (1984) axiom of mere addition expresses a requirement akin to avoidance of the sadistic conclusion. It requires that adding people with positive utilities to any given initial distribution leads to a distribution that is at least as good as the original.

Mere addition. For all $u \in \Omega$ and for all $v \in \Omega_{++}$, $(u, v)Ru$.

As is well-known, Sider's principles satisfy mere addition. This observation raises the question whether avoidance of the sadistic conclusion can be replaced with mere addition in the statement of Theorem 1. The answer is in the affirmative, as stated in the following theorem. Because the observation is closely related to an impossibility result of Blackorby, Bossert, Donaldson, and Fleurbaey (1998, p. 7, Theorem 2), we claim very little credit for it. We provide a proof for completeness but stress that it is, in essence, a mere reformulation of that employed by these authors.

Theorem 2. *Let R be a gains-based variable-value principle that satisfies the limit property. If R avoids the repugnant conclusion, then R violates mere addition.*

Proof. Assume that R is a gains-based variable-value principle that satisfies the limit property and avoidance of the repugnant conclusion. Avoidance of the repugnant conclusion implies that there exist $n \in \mathbb{N}$, $\xi \in \mathbb{R}_{++}$, and $\varepsilon \in (0, \xi)$ such that, for all $m \in \mathbb{N}$ with $m > n$,

$$\xi \mathbf{1}_n R \varepsilon \mathbf{1}_m.$$

Let $\mu \in \mathbb{R}_{++}$ be such that $\mu < \varepsilon$. Because R is a gains-based variable-value principle, this inequality implies that, for all $m \in \mathbb{N}$ with $m > n$,

$$\varepsilon \mathbf{1}_m P \mu \mathbf{1}_m.$$

Let $\delta \in \mathbb{R}_{++}$ be such that $\delta < \mu$. By the limit property, there exists $m > n$ such that

$$\mu \mathbf{1}_m R (\xi \mathbf{1}_n, (\mu - \delta) \mathbf{1}_{m-n}).$$

Because R is transitive, it follows that $\xi \mathbf{1}_n P (\xi \mathbf{1}_n, (\mu - \delta) \mathbf{1}_{m-n})$, contradicting mere addition. ∎

Recent work of Franz and Spears (2020) shows that avoidance of the sadistic conclusion and mere addition become equivalent under some additional axioms including continuity and the Pareto principle; see also Greaves (2017). Theorems 1 and 2 together reassert the similarity between avoidance of the sadistic conclusion and mere addition in the context of the gains-based variable-value principles that satisfy the limit property.

5   Concluding remarks

To the best of our knowledge, our class of gains-based variable-value principles does not appear in the previous literature. It includes a plethora of existing principles as special cases and, thus, our results provide a significant generalization of previously-established impossibility results. As long as the limit property is satisfied, any principle in this class faces the fundamental incompatibility between avoidance of the repugnant conclusion and avoidance of the sadistic conclusion. This complements an earlier observation by Blackorby, Bossert, and Donaldson (2005, p. 164, Theorem 5.4). Blackorby, Bossert, and Donaldson consider principles that use total utility as the criterion to compare any two distributions of the same population size and impose a mild regularity condition that is weaker than the assumption that critical levels exist for all distributions. Their result establishes that these principles cannot simultaneously avoid the repugnant conclusion and the sadistic conclusion. Because all members of their class satisfy the limit property, our impossibility theorem offers a generalization of this earlier observation.

References

[1] Arrhenius, G. (2000). An impossibility theorem for welfarist axiologies. Economics and Philosophy, 16, 247–266.

[2] Arrhenius, G. (forthcoming). Population Ethics: The Challenge of Future Generations. Oxford: Oxford University Press.

[3] Arrhenius, G., J. Ryberg, and T. Tännsjö (2017). The Repugnant Conclusion. In: E.N. Zalta (ed.). The Stanford Encyclopedia of Philosophy (Spring 2017 Edition) ⟨`https://plato.stanford.edu/archives/spr2017/entries/repugnant-conclusion`⟩.

[4] Asheim, G.B. and S. Zuber (2014). Escaping the repugnant conclusion: rank-discounted utilitarianism with variable population. Theoretical Economics, 9, 629–650.

[5] Asheim, G.B. and S. Zuber (forthcoming). Rank-discounting as a resolution to a dilemma in population ethics. In: G. Arrhenius, K. Bykvist, T. Campbell, and E. Finneron-Burns (eds.). Oxford Handbook of Population Ethics. Oxford: Oxford University Press.

[6] Blackorby, C., W. Bossert, and D. Donaldson (2004). Critical-level population principles and the repugnant conclusion. In: J. Royberg and T. Tännsjö (eds.). The Repugnant Conclusion: Essays on Population Ethics, pp. 45–59. Dordrecht: Kluwer Academic Press.

[7] Blackorby, C., W. Bossert, and D. Donaldson (2005). Population Issues in Social Choice Theory, Welfare Economics, and Ethics. New York: Cambridge University Press.

[8] Blackorby, C., W. Bossert, D. Donaldson, and M. Fleurbaey (1998). Critical levels and the (reverse) repugnant conclusion. Journal of Economics, 67, 1–15.

[9] Blackorby, C. and D. Donaldson (1984). Social criteria for evaluating population change. Journal of Public Economics, 25, 13–33.

[10] Bossert, W. (1990). Maximin welfare orderings with variable population size. Social Choice and Welfare, 7, 39–45.

[11] Bossert, W. (forthcoming). Anonymous welfarism, critical-level principles, and the repugnant and sadistic conclusions. In: G. Arrhenius, K. Bykvist, T. Campbell, and E. Finneron-Burns (eds.). Oxford Handbook of Population Ethics. Oxford: Oxford University Press.

[12] Carlson, E. (1998). Mere addition and two trilemmas of population ethics. Economics & Philosophy, 14, 283-306.

[13] Franz, N. and D. Spears (2020). Mere Addition is equivalent to avoiding the Sadistic Conclusion in all plausible variable-population social orderings. Economics Letters, 196, article no. 109547.

[14] Greaves, H. (2017). Population axiology. Philosophy Compass, 12, e12442.

[15] Gustafsson, J.E. (forthcoming). Our intuitive grasp of the repugnant conclusion. In: G. Arrhenius, K. Bykvist, T. Campbell, and E. Finneron-Burns (eds.). Oxford Handbook of Population Ethics. Oxford: Oxford University Press.

[16] Hare, R.M. (1988). Possible people. Bioethics, 2, 279–293.

[17] Huemer, M. (2008). In defence of repugnance. Mind, 117, 899–933.

[18] Hurka, T. (1983). Value and population size. Ethics, 93, 496–507.

[19] Mackie, J.L. (1985). Parfit's population paradox. In: J. Mackie and P. Mackie (eds.). Persons and Values, pp. 242–248. Oxford: Oxford University Press.

[20] Ng, Y.-K. (1986). Social criteria for evaluating population change: an alternative to the Blackorby–Donaldson criterion. Journal of Public Economics, 29, 375–381.

[21] Ng, Y.-K. (1989). What should we do about future generations? Impossibility of Parfit's theory X. Economics & Philosophy, 5, 235–253.

[22] Parfit, D. (1976). On doing the best for our children. In: M.D. Bayles (ed.). Ethics and Population, pp. 100–102. Cambridge, MA: Schenkman.

[23] Parfit, D. (1982). Future generations, further problems. Philosophy and Public Affairs, 11, 113–172.

[24] Parfit, D. (1984). Reasons and Persons. Oxford: Oxford University Press.

[25] Pivato, M. (2020). Rank-additive population ethics. Economic Theory, 69, 861–918.

[26] Sider, T.R. (1991). Might theory X be a theory of diminishing marginal value? Analysis, 51, 265–271.

[27] Spears, D. and M. Budolfson (2021). Repugnant conclusions. Social Choice and Welfare, 57, 567–588.

[28] Tännsjö, T. (2002). Why we ought to accept the repugnant conclusion. Utilitas, 14, 339–359.

[29] Zuber, S. (2018). Population-adjusted egalitarianism. CES Working Paper, halshs-01937766.