Social Choice and Welfare
https://doi.org/10.1007/s00355-021-01321-2

ORIGINAL PAPER

Repugnant conclusions

Dean Spears[1,2,3,4,5] · Mark Budolfson[6,7,8]

Received: 15 July 2020 / Accepted: 9 March 2021
© Springer-Verlag GmbH Germany, part of Springer Nature 2021, corrected publication 2021

Abstract

The population ethics literature has long focused on attempts to avoid the repugnant conclusion. We show that a large set of social orderings that are conventionally understood to escape the repugnant conclusion do not in fact avoid it in all instances. As we demonstrate, prior results depend on formal definitions of the repugnant conclusion that exclude some repugnant cases, for reasons inessential to any "repugnance" (or other meaningful normative properties) of the repugnant conclusion. In particular, the literature traditionally formalizes the repugnant conclusion to exclude cases that include an unaffected sub-population. We relax this normatively irrelevant exclusion, and others. Using several more inclusive formalizations of the repugnant conclusion, we then prove that any plausible social ordering implies some instance of the repugnant conclusion. This understanding—that it is impossible to avoid all instances of the repugnant conclusion—is broader than the traditional understanding in the literature that the repugnant conclusion can only be escaped at unappealing theoretical costs. Therefore, the repugnant conclusion provides no methodological guidance for theory or policy-making, because it does not discriminate among candidate social orderings. So escaping the repugnant conclusion should not be a core goal of the population ethics literature.

1 Introduction

An enduring puzzle in the economics of social welfare is how to incorporate variable population size into social orderings. How should social welfare functions evaluate policies, such as climate policy (Broome 2012; Scovronick et al. 2017), national

This paper received no specific funding, Spears' research is supported by grant K01HD098313 and by P2CHD042849, Population Research Center, awarded to the Population Research Center at The University of Texas at Austin by the Eunice Kennedy Shriver National Institute of Child Health and Human Development. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

✉ Dean Spears
dspears@utexas.edu
http://www.riceinstitute.org

Extended author information available on the last page of the article

Published online: 30 March 2021

health insurance programs, or education subsidies, that will change both the *well-being* and the *number* of future people? Because many interventions will influence population size, this is an important question for economic policy.

It is widely agreed that the population ethics literature is far from fulfilling the goal of providing guidance for these important policy questions. This is because the population ethics literature has long remained focused on the attempt to avoid a condition called the "repugnant conclusion."[1] The repugnant conclusion is an implication of social orderings that allow the quantity of people to compensate for changes in per-person quality of life. Parfit (1984) originally formulated the repugnant conclusion as the hypothetical possibility of a large enough number of lives such that the large number of lives at a low, positive level of utility would be socially preferable to a smaller number of excellent lives, according to some social ordering.

Quantity–quality tradeoffs are at the core of population economics. For every social ordering, there are cases where small changes for some people are socially valued above large changes for others. And yet, the repugnant conclusion has been interpreted as a special implication of only *some* social orderings, such as total utilitarianism (which ranks population according to the sum of wellbeing). Average utilitarianism, for example, (which ranks populations by average wellbeing) is canonically interpreted not to imply the repugnant conclusion.

Our study clarifies the scope and implications of the repugnant conclusion. The leading scholarship in population economics has focused on theorems about which families of social welfare functions do or do not imply the repugnant conclusion (Ng 1989; Blackorby et al. 1998a; Arrhenius 2000; Blackorby et al. 2005; Fleurbaey and Zuber 2015). These impossibility theorems have built an understanding that a social ordering *can* avoid the repugnant conclusion, but only by having some other, worse implication.[2]

This paper contributes formal results and examples that reveal that this understanding of the repugnant conclusion should be revised. "Repugnance" is more common than previously believed. The repugnant conclusion is not an implication of merely *some* social orderings. In fact, under the understanding of the repugnant conclusion that we present, some instance of a repugnant conclusion is an implication of *every* candidate social ordering in population economics.

To reach this conclusion, we show that the literature has used several formal definitions of the repugnant conclusion. Because of path-dependence from prominent original examples, each of these formalizations captures only a subset of

[1] For example, the Stanford Encyclopedia of Philosophy entry on the repugnant conclusion ends its introductory paragraph with "Thus, the question as to how the Repugnant Conclusion should be dealt with and, more generally, what it shows about the nature of ethics has turned the conclusion into one of the cardinal challenges of modern ethics" (Arrhenius et al. 2017).

[2] For example, Ng (1989) proves that any plausible social ordering must *either* imply the repugnant conclusion *or* violate one of two other conditions called Non-Antiegalitarianism and Mere Addition. Similarly, Asheim and Zuber (2014) prove that a family of social orderings "either leads to the Weak Repugnant Conclusion or violates the Weak Non-Sadism Condition." Other important recent examples are the core contributions of Arrhenius (Population ethics: the challege of future generations (unpublished)) and Bossert (2017) and results in McCarthy et al. (2020).

equivalently "repugnant" cases. We show that whether a social welfare function is aggregative or non-aggregative, it must imply one or more new formalizations of repugnant conclusions. Because our new formalizations are only slight modifications of those in the literature, these new formalizations preserve any "repugnance" in traditional formalizations.

The implication is that population ethics is not fundamentally a choice between, on the one hand, ever entailing the repugnant conclusion in some hypothetical choice, or, on the other hand, accepting other undesirable implications. The illusion that this choice exists arises only because of a history of formal definitions that exclude many instances of alleged "repugnance." Similarly, such "repugnance" is not a defining property of total utilitarianism, as it has been previously cast in the literature: for example, there are choice sets from which average utilitarianism implies "repugnant" directives to choose worse lives rather than better lives, but total utilitarianism does not.[3] We conclude that because "repugnance" cannot be escaped, whatever it may be, escaping it should not be a goal. So population ethics should not impose a requirement to avoid the repugnant conclusion.

1.1 Outline

Section 2 introduces our setting and four basic axioms. Section 3 introduces a fundamental question of this paper: What counts as a "repugnant conclusion"? How should Parfit's original example be translated into a family of formal definitions—without begging the question of whether it is bad? We observe that, although the prior literature includes several distinct "repugnant conclusions," purporting to capture the same normative claim, it has restricted formal definitions of the repugnant conclusion to the strict subset of cases where the binary choice between populations includes no intersecting subset of unaffected lives—lives that could be distant in time or space, or in the past. However, *every* policy choice includes unaffected sub-populations, including, at a minimum, the set of people who have already died (Blackorby et al. 1995).

So we define *unrestricted* versions of the repugnant conclusion, which may or may not include such unaffected lives. These new formalizations are what we use in our main result. Later, we define an extended repugnant conclusion, which reflects the fact that any "repugnance" in a quantity–quality tradeoff is as available in fixed-population, same-number cases as it is in the different-number cases of population ethics. If so, the repugnant conclusion is not specific to population ethics.

Section 4 introduces an Aggregation axiom and uses it to present the Theorem that is our main result. All social orderings that satisfy the Aggregation axiom (and one of several specifications of a socially-neutral life) imply a repugnant conclusion, in our unrestricted version. Because Aggregation is closely related to separability in same-number, risk-free cases, this result includes all social orderings that

[3] Although without the same formalization, emphasis, or scope as our paper, prior arguments in this direction have been made in the philosophy literature by Anglin (1977) and Cowen (1996).

economists use, in practice, for policy analysis and many that are commonly understood to escape repugnant conclusions.

But some social orderings deny Aggregation or are otherwise excluded from the main result of our paper. Section 5 discusses examples, while noting further ways in which the formalization of "repugnance" has varied in the literature. So Sect. 6 defines a set of all plausible social orderings—more inclusive than the Aggregative set—and uses this inclusive set to expand our main result to a supporting proposition. In particular, we show that all plausible social orderings that satisfy an axiom called Minimal Equality Preference, whether Aggregative or not, imply an extended repugnant conclusion. Either in aggregating large quantities of tiny changes, or in ignoring them, every social ordering has unintuitive consequences over an unbounded domain. This observation is not unique to population ethics (Cowen 1996; Fleurbaey and Tungodden 2010) and is certainly not specific to totalism and related social orderings. Therefore, the mere fact that a social ordering entails a "repugnant" quantity–quality tradeoff for some example that can be constructed in unbounded space is not informative and cannot guide population ethics.

2 Setting

We largely use the same notation for welfarist, variable-population social evaluation as Blackorby et al. (2005). $\mathbb{Z}$ is the integers, $\mathbb{R}$ is the real numbers, $\mathbb{R}_{++}$ and $\mathbb{R}_{+}$ are the positive and nonnegative real numbers, respectively, and similarly for $-$, $--$, and $\mathbb{Z}$.

Populations $\mathbf{u}, \mathbf{v}$ are finite-length vectors of real numbers. The size of $\mathbf{u}$ is $n(\mathbf{u}) \in \mathbb{Z}_{++}$, so $\mathbf{u} \in \mathbb{R}^{n(\mathbf{u})}$. Population vectors list lifetime utilities or well-beings (terms we will use interchangeably), $\mathbf{u} = (u_1, \dots, u_i, \dots u_{n(\mathbf{u})})$, among the $n(\mathbf{u})$ individuals who make up the population. Because we assume anonymity throughout, the binary relations we consider do not depend on the identities of the individuals; where we informally refer to a "person" we mean only a lifetime well-being in a population. Utilities are normalized so 0 is a neutral lifetime utility.[4] As we discuss in Sect. 4.1, however, adding a life at zero may not be neutral for social evaluation. Following (Asheim and Zuber 2014), an index enclosed in square brackets indicates a rank from worst-off, so $u_{[3]}$ is the utility of the third-worst-off person in $\mathbf{u}$; otherwise indices i do not imply rank.

In comparing the utilities in same-sized populations, $\mathbf{u} \geq \mathbf{v}$ means that $u_i \geq v_i$ for all i; $\mathbf{u} > \mathbf{v}$ means that $u_i \geq v_i$ and $u_i \neq v_i$ for some i; $\mathbf{u} \gg \mathbf{v}$ means $u_i > v_i$ for all i. $\mathbf{1}_n$ is an n-dimensional vector of ones, so $\xi \mathbf{1}_n$ is a population in which all n people have equal utility ξ. For any two populations $\mathbf{u}$ and $\mathbf{v}$, let $(\mathbf{u}, \mathbf{v})$ denote the combined population, so $n((\mathbf{u}, \mathbf{v})) = n(\mathbf{u}) + n(\mathbf{v})$.

The set of all conceptually possible populations is $\Omega = \bigcup_{n \in \mathbb{Z}_{++}} \mathbb{R}^n$. The task in this paper is to describe $\succsim$, which is a social ordering on Ω. $\succsim$ is a binary relation

[4] If lifetime utility is hedonistic, a neutral lifetime utility would be a life that is as good as a life at a limiting case with no experiences.

with the interpretation that $\mathbf{u} \succsim \mathbf{v}$ means that $\mathbf{u}$ is at least as good as $\mathbf{v}$. The asymmetric and symmetric parts of $\succsim$ are $\succ$ and $\sim$, respectively.

Some parts of this paper will use social welfare functions. There, $g : \mathbb{R} \to \mathbb{R}$ is a continuous, increasing, and linear or strictly concave function such that $g(0) = 0$; the purpose of g is to give utility a prioritarian transformation. Also $f : \mathbb{Z}_+ \to \mathbb{R}_+$ is an increasing function such that $f(0) = 0$; the purpose of f is to give a variable-value transformation of population size. In some cases where f and g are identity functions, they are omitted for clarity.

2.1 Basic axioms

We begin with a set of basic axioms on $\succsim$.

Axiom 1 (*Social order*) The relation $\succsim$ is complete, transitive, and reflexive on Ω.

Axiom 2 (*Anonymity*) For all $\mathbf{u}, \mathbf{v} \in \Omega$ such that $n(\mathbf{u}) = n(\mathbf{v})$, if there exists a bijection $\rho : \{1, \dots, n(\mathbf{u})\} \to \{1, \dots, n(\mathbf{u})\}$ such that $u_i = v_{\rho(i)}$ for all i, then $\mathbf{u} \sim \mathbf{v}$.

Axiom 3 (*Continuity*) For all $n, m \in \mathbb{Z}_{++}$ and for all $\mathbf{u} \in \mathbb{R}^n$, the sets $\{\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \succsim \mathbf{u}\}$ and $\{\mathbf{v} \in \mathbb{R}^m : \mathbf{v} \precsim \mathbf{u}\}$ are closed in $\mathbb{R}^m$.

Axiom 4 (*Same-number Pareto*) For all $\mathbf{u}, \mathbf{v} \in \Omega$ such that $n(\mathbf{u}) = n(\mathbf{v})$, if $\mathbf{u} > \mathbf{v}$ then $\mathbf{u} \succ \mathbf{v}$.

In combination, these four axioms imply that a population-sensitive social ordering can be represented as a social welfare function with two arguments: population size and the same-number equally-distributed equivalent. The equally-distributed equivalent (EDE) of a population $\mathbf{u}$, written as $\Xi(\mathbf{u})$, is the utility level that, if given to every member of the population, would result in an equally-ranked (in the sense of $\sim$) same-size population.

Lemma 1 (Blackorby et al. 2005 Theorem 5.2) *Axioms 1–4 are sufficient for there to exist a reduced-form social welfare function $W : \mathbb{Z}_{++} \times \mathbb{R} \to \mathbb{R}$, such that for all $\mathbf{u}, \mathbf{v} \in \Omega$,*

$$\mathbf{u} \succsim \mathbf{v} \Leftrightarrow W(n(\mathbf{u}), \Xi(\mathbf{u})) \geq W(n(\mathbf{v}), \Xi(\mathbf{v})),$$

where W is continuous and increasing in its second argument, and the EDE Ξ has the properties that it is continuous within $\mathbb{R}^n$ for all $n \in \mathbb{Z}_{++}$, that $\Xi(\xi \mathbf{1}_n) = \xi$ for all $n \in \mathbb{Z}_{++}$ and all $\xi \in \mathbb{R}$, and that $\Xi(\mathbf{u})$ is within the closed $\mathbb{R}$-interval bounded by the best and worst-off people in $\mathbf{u}$.

Although the philosophical literature on population ethics contains papers that explore denying each of these axioms,[5] these have been uncontroversial in the economics literature since (Blackorby and Donaldson 1984), and we adopt them throughout this paper.

2.2 Prominent social welfare functions

Several population-sensitive social welfare functions have been named and studied in the literature. One classic approach is Total Utilitarianism, which ranks populations by the sum of utility:

$$V^{TU}(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} u_i. \tag{TU}$$

Closely related is Total Prioritarianism, which ranks populations by the sum of utility, after a concave, increasing transformation:

$$V^{TP}(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} g\left(u_i\right). \tag{TP}$$

TU and TP both are understood to entail the repugnant conclusion, and they are commonly contrasted with alternative functions that are traditionally understood to escape repugnance. One of these is Average Utilitarianism:

$$V^{AU}(\mathbf{u}) = \frac{1}{n(\mathbf{u})} \sum_{i=1}^{n(\mathbf{u})} u_i. \tag{AU}$$

Another, which is commonly described as a "compromise" that behaves like TU for smaller populations and AU for larger populations, is Variable-Value or Number-Dampened Utilitarianism (Hurka 1983; Ng 1989):

$$V^{NDU}(\mathbf{u}) = \frac{f(n(\mathbf{u}))}{n(\mathbf{u})} \sum_{i=1}^{n(\mathbf{u})} u_i, \tag{NDU}$$

where f is increasing, concave, and bounded.

Since its proposal by Blackorby and Donaldson (1984), a focus of the population economics literature has been Critical-Level Generalized Utilitarianism. CLGU is an additively-separable sum of transformed utility, where the transformation includes a positive critical level $c > 0$ for adding a positive life to constitute a social improvement:

5 In the philosophical literature, for example, Temkin Larry (2014) considers denial of the transitive part of Axiom 1, Roberts (2011) denies Axiom 2, and Carlson (On some impossibility theorems in population ethics (forthcoming)) denies Axiom 3. Such issues are a focus of a companion working paper in the philosophy literature (Budolfson and Spears 2018), which does not contain the formal results of this paper.

$$V^{CLGU}(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} \left[g\left(u_i\right) - g(c)\right]. \tag{CLGU}$$

A recent advancement of this literature is Rank-Discounted Generalized Utilitarianism (RDGU). RDGU was proposed by Asheim and Zuber (2014) and further investigated by Pivato (2020). This social welfare function transforms each person's utilities u_i by an increasing function g and weights utilities by a weight that is geometrically decreasing in rank-distance from the worst-off person:

$$V^{RDGU}(\mathbf{u}) = \sum_{r=1}^{n(\mathbf{u})} \beta^r g\left(u_{[r]}\right), \tag{RDGU}$$

where $0 < \beta < 1$ and the square-bracket index $[r]$ indicates that the population $\mathbf{u}$ is ordered in increasing rank. As Asheim and Zuber prove, RDGU escapes Parfit's formulation of the repugnant conclusion.

3 Repugnant conclusions

Whether the repugnant conclusion can be avoided depends on what the repugnant conclusion is. Our goal in this section is to avoid begging the question of whether the repugnant conclusion is bad or which social orderings entail it. Because Parfit's (1984) original statement of the repugnant conclusion invokes a well-off population of ten billion people, every paper in the formal population literature has adopted a formal definition that goes beyond Parfit's original example. These formalizations sometimes disagree, and even (Parfit 2016) has written about heterogeneity among instances of repugnant conclusions.[6] Yet the population economics literature typically formalizes the repugnant conclusion as:

Definition 1 (*The (original, restricted) repugnant conclusion*) For any $u^h \in \mathbb{R}_{++}$, any $n^h \in \mathbb{Z}_{++}$, and any $\varepsilon > 0$, there exists $n^\varepsilon \in \mathbb{Z}_{++}$ such that $\varepsilon \mathbf{1}_{n^\varepsilon} \succ u^h \mathbf{1}_{n^h}$.

In Definition 1, the populations $\varepsilon \mathbf{1}_{n^\varepsilon}$ and $u^h \mathbf{1}_{n^h}$ do not overlap: there is no intersecting utility level v_j of person j who lives the same life, irrespective of whether $\varepsilon \mathbf{1}_{n^\varepsilon}$ or $u^h \mathbf{1}_{n^h}$ is chosen. But, as Parfit (1984) also noted, "these questions [of population ethics] arise most clearly when we compare the outcomes that would be produced, in the further future, by different rates of population growth." Any policy choice that changes the *future* leaves the *past* unaffected. So the full consequences of any actual policy choice include many lives that intersect, unchanged in both possible populations: past lives, at least, and plausibly more, as well. A central insight of Blackorby et al. (1995) "independence of the utilities of the dead" axiom is that populations

[6] Parfit (2016) writes of 'a,' 'another,' 'this,' and 'a version of the' repugnant conclusion.

Table 1 Repugnant conclusions

	$n^{\ell} = 0$	$n^{\ell} \geq 0$
$n(\mathbf{v}) = 0$	Restricted repugnant conclusion	Restricted very repugnant conclusion
$n(\mathbf{v}) \geq 0$	Repugnant conclusion	Very repugnant conclusion

exist in time, and past populations cannot be influenced by future choices.[7] The existence of a dead, past subpopulation—or of any other unaffected population—is irrelevant to any repugnance in the choice to create n^{ε} lives at ε rather than n^h lives at u^h. So there is no normative reason to restrict the repugnant conclusion to cases without an unaffected, intersecting population. Therefore, in Definition 2 we redefine the repugnant conclusion to remove this normatively irrelevant restriction and to permit an intersecting sub-population $\mathbf{v}$, which may be empty, and may live in a distant time or place:

Definition 2 (*The (unrestricted) repugnant conclusion*) For any $u^h \in \mathbb{R}_{++}$, any $n^h \in \mathbb{Z}_{++}$, and any $\varepsilon > 0$, there exist $n^{\varepsilon} \in \mathbb{Z}_{++}$ and $\mathbf{v} \in \Omega \cup \{\emptyset\}$ such that $\left(\varepsilon \mathbf{1}_{n^{\varepsilon}}, \mathbf{v}\right) \succ \left(u^h \mathbf{1}_{n^h}, \mathbf{v}\right)$.

Thus, our contribution begins by offering a revised formalization of the repugnant conclusion. In our terminology, Definition 2 is *the* repugnant conclusion.[8] Definition 1 is the *restricted* repugnant conclusion.

Arrhenius (2003) introduced the *very* repugnant conclusion, which intensified the repugnant conclusion by stipulating that the ε-lives are accompanied by a large number of highly *negative* utility lives, full of suffering and not nearly worth living, which could be avoided by choosing the u^h-lives. Like Parfit's original example of the repugnant conclusion, Arrhenius' very repugnant conclusion is restricted, in our formal sense: it does not include an intersecting, unaffected subpopulation $\mathbf{v}$. Therefore, we introduce an unrestricted definition, which we propose should be used to capture further instances of the very repugnant conclusion:

Definition 3 (*The (unrestricted) very repugnant conclusion*) For any $u^h \in \mathbb{R}_{++}$, any $n^h \in \mathbb{Z}_{++}$, any $u^{\ell} \in \mathbb{R}_{--}$, any $n^{\ell} \in \mathbb{Z}_{++}$, and any $\varepsilon > 0$, there exist $n^{\varepsilon} \in \mathbb{Z}_{++}$ and $\mathbf{v} \in \Omega \cup \{\emptyset\}$ such that $\left(\varepsilon \mathbf{1}_{n^{\varepsilon}}, u^{\ell} \mathbf{1}_{n^{\ell}}, \mathbf{v}\right) \succ \left(u^h \mathbf{1}_{n^h}, \mathbf{v}\right)$.

[7] Blackorby et al. (1995) use this observation to motivate an additively separable approach to population ethics, on the grounds that the utilities of past people should not influence the evaluation of policies that only impact future people; although our results include non-separable social orderings such as average utilitarianism, we further discuss separability in Sect. 4.3.

[8] Arrhenius (Population ethics: the challege of future generations (unpublished)) uses an equivalent condition called the Strong Quality Addition Principle; Anglin (1977) shows that this principle is implied by both total and average utilitarianism, and Arrhenius extends this proof to Ng's (1989) variable-value utilitarianism. Note that our theorem below uses a different condition, the unrestricted very repugnant conclusion.

To our knowledge no equivalent condition has previously appeared in the population ethics literature under any name. Wherever we refer to the "very repugnant conclusion" below, we mean the unrestricted version in Definition 3. The restrictions and subsets can be understood and compared using Table 1, which summarizes this paper's revision and generalization of terminology in the literature for repugnant conclusions.

To emphasize, we do not presume that any version of these conclusions is in fact normatively repugnant. In constructing Table 1 and this section's definitions, we have not answered the question of what "repugnance" fundamentally is or may be, if any, in any quantity–quality population tradeoff. Instead, our concluding arguments will informally assume a conditional claim about normative repugnance:

Conditional claim about repugnance If the fact that a social order $\succsim$ implies a restricted repugnant conclusion with $n(\mathbf{v}) = 0$ is sufficient to disqualify $\succsim$ normatively, then any $\succsim'$ that implies the unrestricted $n(\mathbf{v}) \geq 0$ equivalent of an otherwise identical repugnant conclusion is similarly disqualified.

In our conclusion, we will argue that because unrestricted repugnant conclusions are too ubiquitous to be disqualifying, restricted repugnant conclusions must not be disqualifying, either, according to the contrapositive of our conditional claim. In support of our conditional claim, we emphasize that the restriction that $n(\mathbf{v}) = 0$ is irrelevant to the repugnance that some perceive in a social ordering choosing arbitrarily many arbitrarily negative lives, along with a large number of barely-positive lives, when arbitrarily many arbitrarily wonderful lives were possible instead. More formally, if one assumes independence of the utilities of the dead—or existence independence[9] more generally—then any restricted repugnant conclusion is logically equivalent to its unrestricted version, although one need not accept any form of independence to accept our conditional claim about repugnance.

Ultimately, if the goal is to choose among actual population and economic policies, then these can only influence the future. So for any actual policy choice, $n(\mathbf{v}) > 0$. Noticing this, Dasgupta (2005) labels hypothetical choices where $n(\mathbf{v}) = 0$ as "Genesis problems," and dismisses them as "the wrong problem."[10] In proceeding with unrestricted repugnant conclusions, we do not follow Dasgputa in ignoring cases where $n(\mathbf{v}) = 0$, but nor do we ignore cases where $n(\mathbf{v}) > 0$. Therefore, we conclude that the cases where $n(\mathbf{v}) > 0$ are at least as normatively and practically important as the restricted cases where $n(\mathbf{v}) = 0$. There is no normative or practical reason to impose a constraint to the restricted subset.

[9] The existence independence axiom holds that for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \Omega$, $(\mathbf{u}, \mathbf{w}) \succsim (\mathbf{v}, \mathbf{w})$ if and only if $\mathbf{u} \succsim \mathbf{v}$ (Blackorby et al. 1995, p. 159).

[10] Dasgupta (2005) elaborates: "The Genesis Problem may have been God's problem, but it is not the problem we face. We are here."

4 Main result: aggregation and its consequences

This paper partitions the social orderings defended in the population economics literature into those that satisfy an axiom of Aggregation and those that do not. This section focuses on the former; Sect. 6 focuses on the latter. Here, Sect. 4.1 introduces axioms which interpret the neutral, zero level of utility (used in the repugnant conclusion's emphasis on $\varepsilon > 0$ lives). Then, Sect. 4.2 presents the Aggregation axiom and, with it, the Theorem which is our main result. Finally, Sect. 4.3 discusses properties that are sufficient for a social ordering to satisfy Aggregation.

4.1 The sign of lifetime utility

The repugnant conclusion invokes lives at $\varepsilon > 0$. Such slightly-positive wellbeing is how the population economics literature has long formalized Parfit's phrase "lives that are barely worth living." None of the basic axioms have yet distinguished among lives that are 0, positive, or negative. For Parfit's original repugnant conclusion to be meaningful, we must make an assumption about these lives. Indeed, if there is no meaningful or obvious assumption to be made about lives at and above zero, it is not clear why any conclusion about them would be "repugnant."

What are the formal implications of a life that is neutral? As Blackorby et al. (2005) summarize, "a person who lives a neutral life considers this life to be as good as a life without any experiences" (p. 161), but this fact need not imply that the *social evaluation* considers adding such a life not to change the value of the full population. The classic zero axiom, named "mere addition" by Parfit, is the weaker claim that adding a life of positive utility does not make a population worse:

Definition 4 (*Mere addition*) For all $\mathbf{v} \in \Omega$, $u \in \mathbb{R}_{++}$, it is the case that $\left(\mathbf{v}, u\mathbf{1}_1\right) \succsim \mathbf{v}$.

Many social orderings in the literature do not satisfy mere addition (Blackorby et al. 1998a; Franz and Spears 2020). Average utilitarianism and Ng's (1989) variable-value utilitarianism fail mere addition, because additional positive lives could lower average utility. So does RDGU (Pivato 2020).

Therefore, instead of mere addition, our Theorem allows a social ordering to satisfy either of two arguably more attractive sign axioms. Average utilitarianism, variable-value utilitarianism, and RDGU each satisfy Axioms 5 and 6, as do orderings such as maximin and maximax with reduced forms that are insensitive to n. CLGU, however, satisfies neither sign axiom. Axioms 5 and 6 only apply to perfectly equal populations, to avoid the problematic cases for average-type theories where additional positive lives bring down the population-wide average.

Axiom 5 (*First sign axiom: egalitarian dominance*) For any $u, v \in \mathbb{R}_{++}$ such that $u > v > 0$ and any $n, m \in \mathbb{Z}_{++}$ such that $n > m$, we have that $u\mathbf{1}_n \succ v\mathbf{1}_m$.

Axiom 6 (*Second sign axiom: egalitarian priority for lives worth living*) For any $u, v \in \mathbb{R}$ such that $u > 0 > v$ and any $n, m \in \mathbb{Z}_{++}$, we have that $u\mathbf{1}_n \succ v\mathbf{1}_m$.

4.2 Axiom of aggregation

The last step before our main result is an axiom of Aggregation. We interpret it to reflect a weak commitment to not-so-unequal consideration of the interests of the full set of people who ever exist. Although formally distinct from the aggregation axiom of Fleurbaey and Tungodden (2010) (see details in Sect. 6), it reflects the same intuition: that a bounded loss or gain accruing to only a small part of the population cannot have a large effect on the social evaluation if the consequences are different for everybody else.

Axiom 7 (*Aggregation*) For any $\mathbf{u} \in \Omega$, any $\delta \in \mathbb{R}_{++}$, and any utility level $\xi \in \mathbb{R}$, there exists $n^* \in \mathbb{Z}_{++}$ such that if $n \geq n^*$, then:

$$(\xi + \delta)\mathbf{1}_{n+n(\mathbf{u})} \succ \left(\xi\mathbf{1}_n, \mathbf{u}\right) \succ (\xi - \delta)\mathbf{1}_{n+n(\mathbf{u})}.$$

The Aggregation axiom holds that the equally-distributed equivalent becomes diminishingly sensitive to any consequence for a small subset of the population, as that subset becomes a small enough part of a large enough population. Note that the comparison in the axiom's consequent clause holds population size constant. As Sect. 4.3 details, many social orderings in the population economics literature satisfy Aggregation.

Theorem 1 *If* $\succsim$

- *is a social order (Axiom 1) that satisfies anonymity (Axiom 2),*
- *satisfies Aggregation (7), and*
- *satisfies at least one of the sign Axioms (5 or 6),*

then $\succsim$ *implies the very repugnant conclusion (moreover, if Axiom 5 is satisfied, the unaffected population* $\mathbf{v}$ *in the very repugnant conclusion can be restricted to be positive, so* $\mathbf{v} \gg \mathbf{0}$*).*

Proof See Appendix A.1. □

So the aggregative social orderings that satisfy the conditions of Theorem 1 all imply repugnant conclusions, even though they take radically different approaches to different-number population ethics. The cases that are repugnant for one social ordering may not be for another. An important observation is that the binary choices in which, for example, total utilitarianism would make a repugnant choice are not a superset of the choices in which average utilitarianism would make a repugnant choice.[11] So the repugnant conclusion offers no argument against additively

[11] To see this, consider a case where $\varepsilon = 10^{-6}$, $n^\varepsilon = 10{,}000$, $u^h = 9$, $n^h = 10$, and $\mathbf{v}$ is 10 lives at -10 (for simplicity, let $n^\ell = 0$). Average utilitarianism would choose the ε lives and total utilitarianism would choose the u^h lives. If $\varepsilon = 0$, which would only increase any normative repugnance of such choice, average utilitarianism would continue to choose the ε lives as n^ε becomes ever larger, but total utilitarianism would continue to choose the u^h lives.

separable social orderings, in particular, in favor of aggregative but non-separable alternatives. Note that Continuity[12] (Axiom 4) and Pareto (Axiom 3) are not needed; Anonymity is used only in that it is implicitly assumed in our definitions throughout.

4.3 Which social welfare functions satisfy aggregation?

The significance of the Theorem is in the contrast between the extent of the Theorem's scope, on the one hand, and the conventional wisdom about the repugnant conclusion, on the other hand. The population economics literature contains many studies which contrast average and total utilitarianism (sometimes called Millian and Benthemite social welfare functions, in this literature) as alleged opposite approaches to social evaluation (*e.g.*, Nerlove et al. 1982).[13] On this common view, the repugnant conclusion is widely understood to be a problematic implication only of total utilitarianism, total prioritarianism, and related totalist social objectives. This understanding allegedly offers a reason to reject these orderings in favor of alternatives such as average utilitarianism or variable-value utilitarianism, which would not imply the repugnant conclusion. The Theorem tells us that this is a misunderstanding, because all four of these social orderings (and more) imply the very repugnant conclusion, properly understood.

A sufficient condition to satisfy Aggregation is for the reduced-form representation to take the separable form, which satisfies the four basic axioms:

$$V(\mathbf{u}) = W\left(n(\mathbf{u}), \sum_{i=1}^{n(\mathbf{u})} g\left(u_i\right) \right) \tag{1}$$

This family of functional forms is common in the population economics literature. It includes total, average, variable-value, and critical-level versions of utilitarianism, prioritarianism, and egalitarianism.[14] Deschamps and Gevers (1977) summarize a standard view in welfare economics: "the separability axiom seems to have considerable appeal ⋯ when an ethical observer engages in social welfare judgements which involve no uncertainty."

One reason that the family in equation 1 is attractive is because it satisfies same-number independence:

[12] Gustafsson (2017) offers an example of a social ordering that satisfies Aggregation but not Continuity.

[13] Further examples include (Palivos and Yip 1993; Dasgupta 2005; Boucekkine and Fabbri 2013; Spears 2017; Scovronick et al. 2017), and Lawson and Spears (2018).

[14] In this paper we distinguish between prioritarianism and egalitarianism using the definitions of Broome (2015). Both functional forms use concave g transformations and same-number risk-free additive separability, but the prioritarian social welfare function is additively separable, while egalitarianism follows (Fleurbaey 2010) in inverting g to use the EDE, so average egalitarianism is $W^{AE}(\mathbf{u}) = g^{-1}\left(\frac{1}{n(\mathbf{u})} \sum_{i=1}^{n(\mathbf{u})} g(u_i)\right)$ and total egalitarianism is $W^{TE}(\mathbf{u}) = n(\mathbf{u})g^{-1}\left(\frac{1}{n(\mathbf{u})} \sum_{i=1}^{n(\mathbf{u})} g(u_i)\right)$. Total prioritarianism satisfies mere addition but total egalitarianism does not; both satisfy the conditions of the Theorem and therefore imply the very repugnant conclusion. Nothing hinges on our use of this terminology from the literature, however.

Axiom 8 (*Same-number independence*) For any $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x} \in \Omega$ such that $n(\mathbf{u}) = n(\mathbf{v})$ and $n(\mathbf{w}) = n(\mathbf{x})$, if $(\mathbf{u}, \mathbf{w}) \succsim (\mathbf{v}, \mathbf{w})$ then $(\mathbf{u}, \mathbf{x}) \succsim (\mathbf{v}, \mathbf{x})$ and similarly for $\succ$.

Consider the large set of ordinary, non-population economic policy decisions. Should taxes transfer more from the rich to the poor? Should schools invest more in younger or older children? If these policies do not change the size of the population, then they are same-number questions. If the social ordering does not satisfy same-number independence, then assessing these policies requires knowing the utility of unaffected people in distant places and times—which would violate independence of the utilities of the dead.

Same-number independence is not quite sufficient for the functional form in equation 1. As Blackorby et al. (1998b) and Blackorby et al. (2005) show, in the context of the basic Axioms (1–4), same-number independence is sufficient for there to exist a set of population-size-indexed increasing and continuous functions g^n such that the EDE has an additively separable structure for all n:

$$\Xi(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} g^{n(\mathbf{u})}\left(u_i\right). \tag{2}$$

Equation 2, unlike equation 1, permits g to differ by n, which could prevent Ξ from converging as Aggregation requires. A sufficient condition, in the context of the basic axioms, for same-number independence to imply the form in equation 1 is replication invariance (*cf.*Fleurbaey and Tungodden 2010; Asheim and Zuber 2018). However, replication invariance is not necessary for our results. It is sufficient for same-number separability (Axiom 8) and the basic Axioms (1–4) to imply Aggregation if $\frac{g^n(u)}{n}$ goes to 0 as n goes to infinity for all fixed $u \in \mathbb{R}$.

Corollary 1 *If* $\succsim$ *satisfies the basic Axioms (1−4) and same-number independence (8), then it has a same-number separable reduced-form representation (Lemma 1); if additionally it satisfies at least one of the sign Axioms (5 or 6) and its set of* g^n *satisfy that* $\lim_{n\to\infty} \frac{g^n(u)}{n} = 0$ *for all* $u \in \mathbb{R}$, *then* $\succsim$ *implies the very repugnant conclusion.*

To emphasize, Aggregation is a *different-number* axiom of population ethics, but it is a consequence of *same-number* independence (familiar from ordinary, fixed-population economic policy analysis) combined with the different-number condition that $\lim_{n\to\infty} \frac{g^n(u)}{n} = 0$, which would be implausible for a same-number-separable social ordering to reject (indeed, no such social ordering is defended in the literature, to our knowledge).

5 Examples of "repugnance" beyond our main result

What is distinctive about the repugnant conclusion that separates it from all of the other kinds of objections that might be raised to a social ordering? Section 3 argued that our understanding of the repugnant conclusion should include "addition" cases, and that it is inappropriately restricted if it does not. Section 4 demonstrated our main result: that with this formalization of the repugnant conclusion, its reach is much wider than commonly understood. But, as plausible as the set of social orderings covered by Theorem 1 may be, some attractive social evaluations are excluded from this set. For the rest of the paper, we expand our results further, continuing our argument that normatively similar cases should not be excluded from the formalization of the repugnant conclusion.

Maximin, RDGU, and CLGU escape Theorem 1. Neither maximin nor RDGU satisfies Aggregation. CLGU satisfies Aggregation but does not satisfy either of the sign axioms unless it is “standardized” such that the critical level is zero[15]. All of these social orderings are addressed by Proposition 1, below. Here, we briefly note that the understanding in the literature that CLGU and RDGU escape the repugnant conclusion depends on *further* ways in which the formalization of the repugnant conclusion varies in the literature.[16] Each of Blackorby and Donaldson (2005, p. 162), Asheim and Zuber (2014), and Arrhenius (Population ethics: the challege of future generations (unpublished), p. 403) state a formalization of the repugnant conclusion in which, implicitly, $n(\mathbf{v}) = 0$. However, these formalizations differ:

Population size. Blackorby et al. (2005) and Asheim and Zuber (2014) require that $n^{\varepsilon} > n^{h}$, but Arrhenius (Population ethics: the challege of future generations (unpublished)) does not include any requirements about the size of either population. In other words, only some formalizations require that n^{ε} be large. It is surely at least as repugnant to choose a lower quantity of lower-quality lives over a larger quantity of higher-quality lives.

Utility levels. Arrhenius (Population ethics: the challege of future generations (unpublished)) interprets ε-lives qualitatively as “barely worth living.” But Blackorby et al. (2005) and Asheim and Zuber (2014) merely require that they be worse than u^{h}, which is not required to be high; these authors require only that $u^{h} > \varepsilon > 0$.

Because they differ in two ways, neither definition is strictly more inclusive or exclusive than the other. The most inclusive *combination* of these properties would consider any choice of any equal lower-utility population over any equal higher-utility population an instance of the repugnant conclusion, provided that both utility levels were positive, irrespective of population size and of the utility levels. This occurs under both RDGU and CLGU. For both RDGU (for any fixed β) and CLGU

[15] Broome (2004) advances a case for CLGU which he “standardizes” by setting the critical level equal to zero, adjusting g to match. (In this case, zero may or may not be a neutral life for the person.) Broome interprets his resulting social ordering to imply the repugnant conclusion, which he argues is unintuitive but ultimately acceptable. What Broome there calls “the repugnant conclusion,” Arrhenius (Population ethics: the challege of future generations (unpublished)) names “the weak repugnant conclusion,” a further example of simultaneous debate in the prior literate about the *extent* and *acceptability* of the repugnant conclusion.

[16] To our knowledge, we are the first to note these discrepancies or their implications.

(for any fixed c) there exist cases where $x, y \in \mathbb{R}_{++}$, $x > y$, y is "barely worth living," and $n(\mathbf{v}) = 0$ such that $y\mathbf{1}_m \succ x\mathbf{1}_n$ for some $m, n \in \mathbb{Z}_{++}$.[17] So, we may ask: is this, too, a "repugnant" conclusion?

Our purpose in this example is to emphasize that, although social orderings can be constructed that escape some instances or formalizations of "repugnance," such results should not be conflated with escaping *all* instances of a comparably "repugnant" quantity–quality tradeoff. In some cases, indeed, the entailed "repugnance" is even more extreme—such as choosing a *smaller* worse-off population. Motivated by these examples and by the ambiguity in the literature about what "repugnance" is, the next section pursues a broad extension of our main result.

6 Beyond aggregation

In this section we broaden our scope. We present an extended proposition that includes Non-Aggregative social orderings. Denying Aggregation permits gains to the worst-off members of the population to outweigh consequences for the better-off rest of the population. So some evaluators find Non-Aggregative social orderings a reasonable tool to capture normative values. Here we also include orderings that reject the sign axioms.

To include Non-Aggregative social orderings, we weaken Aggregation to create an axiom that includes all plausible candidate social orderings, Minimal Equality Preference:

Axiom 9 (*Minimal Equality Preference*) For any $\mathbf{u} \in \Omega$, any $\delta \in \mathbb{R}_{++}$, and any utility level $\xi \in \mathbb{R}$, there exists $n^* \in \mathbb{Z}_{++}$ such that if $n \geq n^*$, then:

$$(\xi + \delta)\mathbf{1}_{n+n(\mathbf{u})} \succ \left(\xi\mathbf{1}_n, \mathbf{u}\right).$$

Minimal Equality Preference weakens Aggregation only by deleting "$\succ (\xi - \delta)\mathbf{1}_{n+n(\mathbf{u})}$" from the consequent clause. It is named as a reference to the "Weak Equality Preference" axiom of Blackorby et al. (1998a), which implies Minimal Equality Preference in the context of our basic axioms. Where we refer to "Non-Aggregative" social orderings, we mean those that satisfy Minimal Equality Preference but do not satisfy Aggregation.[18] These include maximin, critical-level leximin, and RDGU.

[17] To see this, for RDGU, choose y that is very close to x, let $n = 1$ and let $m > 1$. For CLGU let $c > x > y > 0$ and n be much larger than m (a violation for CLGU of our Axiom 5).

[18] Fleurbaey and Tungodden (2010) prove the incompatibility of their similar Minimal Aggregation and Minimal Non-Aggregation axioms. As the names suggest, their Minimal Aggregation axiom is weaker than our Aggregation axiom. Yet, their Minimal Aggregation and Minimal Non-Aggregation only conflict in the presence of their Reinforcement axiom, which they call a "basic consistency requirement." However, this is not sufficient to cover all candidates for population ethics because RDGU—which was introduced by Asheim and Zuber (2014) after Fleurbaey and Tungodden's (2010) theorem—does not satisfy Reinforcement.

To include all plausible candidate social orderings, we must broaden our formalization of “repugnance.” Again, we can do this with a small change that preserves the repugnance that some perceive in the quantity–quality tradeoffs of welfarist social orderings. The aggregate normative consequences of tiny changes have been explored and sometimes criticized as thoroughly in same-number cases as in population ethics’ different number cases (Cowen 1996; Fleurbaey and Tungodden 2010). To capture this, we define a general, arbitrarily small change to the welfare distribution of a population:

Definition 5 (*ε-change*) Population $\mathbf{u}$ is distinguished only by an ε-change from $\mathbf{v}$ if either:

- $n(\mathbf{u}) = n(\mathbf{v}) + 1$ and $\left(\mathbf{v}, \varepsilon\mathbf{1}_1\right) = \mathbf{u}$, or
- $n(\mathbf{u}) = n(\mathbf{v})$, there is one j such that $u_j = v_j + \varepsilon$, and $u_i = v_i$ for all $i \neq j$.

If a population is distinguished from another by two or more ε-changes, then any one person or potential person in the population may receive at most one ε-change.

Note the “at most one” restriction: although many ε increments could, in principle, add up to a large improvement for one person, our definition only includes cases where each person’s wellbeing changes at most by a tiny amount. People who receive or are created by an ε-change cannot receive a further ε-change, to fulfill this definition. With the definition of ε-change, we can define an extended very repugnant conclusion. The (unrestricted) very repugnant conclusion is the strict subset of cases of the extended very repugnant conclusion in which all ε-changes are additions of $\varepsilon\mathbf{1}_1$.

Definition 6 (*Extended very repugnant conclusion*) For any positive utility level $u^h \in \mathbb{R}_{++}$, any $n^h \in \mathbb{Z}_{++}$, any negative utility level $u^\ell \in \mathbb{R}_{--}$, any $n^\ell \in \mathbb{Z}_{++}$, and any $\varepsilon > 0$, there exist:

- a number of ε-changes $n^\varepsilon \in \mathbb{Z}_{++}$,
- a number of potential negative-utility lives $m^\ell \geq n^\ell$,
- a number of potential positive utility lives $m^h \geq n^h$, and
- an unaffected, intersecting population (which could be empty) $\mathbf{v} \in \Omega \cup \{\emptyset\}$

such that, rather than create the u^h-lives, it is better to instead create the u^ℓ-lives while also having ε-changes, at most one to each person or potential person, *i.e.*:

- there exists $\mathbf{v}^\ell \in \Omega$ which is distinguished only by n^ε ε-changes from $\left(u^\ell \mathbf{1}_{m^\ell}, \mathbf{v}\right)$, and
- $\mathbf{v}^\ell \succ \left(u^h \mathbf{1}_{m^h}, \mathbf{v}\right)$.

The extended very repugnant conclusion—like the original restricted very repugnant conclusion—holds that many terrible lives full of suffering, which need never be lived, should be created, when many wonderful lives are instead available, merely so some other people each receive one tiny benefit. Note that, because ε-changes are tiny, the people who receive one will still have a very negative life after the change, if they did before.

The extended very repugnant conclusion appears awkward and hard to evaluate. But the only difference between the extended very repugnant conclusion and the very repugnant conclusion is that the extended very repugnant conclusion permits ε -changes to accrue to new *or* existing people, while the very repugnant conclusion only permits new people. If quantity–quality tradeoffs among tiny benefits, accessible high-utility lives, and avoidable negative-utility lives are inevitably "repugnant"—if they are axiomatically known to be repugnant prior to theory—then the extended very repugnant conclusion must fit the charge of alleged "repugnance." Yet every plausible social ordering in the literature implies this conclusion:

Proposition 1 *If* $\succsim$

- *satisfies the basic axioms (1-4), and*
- *satisfies either Aggregation (7) or Non-Aggregation [*i.e. *satisfies Minimal Equality Preference],*

then $\succsim$ *implies the extended very repugnant conclusion.*
Proof See Appendix A.2. □

Note that Proposition 1, unlike the Theorem, does not require sign Axioms (5 and 6), because ε-changes need not involve lives near zero utility. The extended very repugnant conclusion is entailed by every social ordering that we are aware to be defended in the population economics literature. It is implied by maximax. It is implied by Sider's (1991) Geometrism. It is implied by odd but imaginable examples such as $V(\mathbf{u}) = \left(\frac{1}{n}\sum_{i=1}^{n(\mathbf{u})} g(u_i)\right) - \frac{\alpha}{n(\mathbf{u})}$, for $\alpha > 0$ (suggested for illustration by Partha Dasgupta).

Of course, we do not actually believe that every social welfare function is normatively repugnant. So we must conclude that the extended very repugnant conclusion is not axiomatically repugnant, after all. But there is little that normatively distinguishes the extended very repugnant conclusion from the "repugnance" of the very repugnant conclusion. With its negative lives, it is arguably more "repugnant" than Parfit's original example. So we conclude that "repugnant" conclusions may not be so repugnant, after all.

7 Conclusion

> "It is time to retire the Repugnant Conclusion from population ethics."
> (Dasgupta 2016)
> "Avoiding the Repugnant Conclusion is not a necessary condition for a minimally adequate candidate axiology, social ordering, or approach to population ethics."

(Zuber et al. 2021)[19]

Our contribution begins in recognizing that the repugnant conclusion, as it has been used in the formal population ethics literature, has been limited to only a subset of equivalently "repugnant" cases. Although the prior literature has formalized the repugnant conclusion in a variety of substantively distinct ways, it has overlooked the importance of an unaffected subset of the population. So we use a more inclusive formalization of the repugnant conclusion.

Our Theorem 1 has a three-part structure that generalizes the three parts of Ng's (1989) impossibility theorem, while retaining or intensifying their normative importance.[20] We weaken logically, but not normatively, the usual formalization of "repugnance." It is not a surprise that such a change has the consequence that more social orderings imply it. Instead, what is important about these results is their *extent*: all social orderings in the population economics literature—and more—imply *some* instance of a repugnant conclusion. Theorem 1 makes this point especially strongly for Aggregative orderings. Totalism is therefore not qualitatively special in this way. The conventional wisdom about the repugnant conclusion merely reflects an arbitrary boundary drawn through a map of similarly repugnant cases.

The implication is that the repugnant conclusion cannot fully be escaped. Therefore, implying "repugnance" offers little methodological guidance in the choice among social welfare functions. In this way, our method and conclusion compare with those of Fleurbaey and Tungodden (2010), who have a related but different substantive focus. They show that all plausible social orderings imply either a Tyranny of Aggregation or a Tyranny of Non-Aggregation, and conclude that "one should be cautious when criticizing maximin, (generalized) utilitarianism or any other social ordering on the basis of how they perform in extreme cases. The assessment of the various possible social ordering functions should be more comprehensive and, maybe, more focused on cases that are directly relevant to actual policy issues." We have argued that, once an arbitrary distinction is removed from its formalization, the repugnant conclusion gives no theoretical guidance. Nor, following Fleurbaey and Tungodden, does the repugnant conclusion give any practical guidance nor apply in feasible cases. So axiomatic avoidance of a repugnant conclusion should be dropped as a methodological requirement for population economics.[21]

[19] This quotation is from a collaboration of 29 authors from economics and philosophy, asking "What should we agree on about the Repugnant Conclusion?". Further related conclusions include those of Adler (2009) ("Perhaps the best solution, on balance, is to revert to 'total' prioritarianism and accept the repugnant conclusion.") and Cowen (2018) (concluding an appendix about the repugnant conclusion by dismissing its relevance: "I say full steam ahead").

[20] Our weak sign axioms play a part similar to mere addition (but only apply to perfectly equal populations); our Aggregation axiom has a role similar to Ng's Non-Antiegalitarianism; and we use an unrestricted repugnant conclusion.

[21] One question is what Parfit himself might have thought of our observation that we misunderstand the lesson of the repugnant conclusion if we impose the normatively irrelevant restriction that $n(\mathbf{v}) = 0$. Late in his career, in a remarkable final paper, Parfit (2017) wrote about a plurality of related conditions: "such repugnant conclusions" (p. 124) he wrote there, just as also in 2016 he wrote about multiple repugnant conclusions (we cite above). In this final paper, although he still sought to escape the repugnant conclusion, Parfit came to a revised understanding of the repugnant conclusion that resonates with ours: "Because the Repugnant Conclusion seemed to me very implausible, I claimed that we ought to reject

Following our conclusion leaves open which family of social orderings to choose. Although disregarding the repugnant conclusion would remove a famous alleged disadvantage of total utilitarianism, other considerations may recommend a different choice. Zuber and Asheim (2012), for example, advocate RDGU without reference to the repugnant conclusion, because its approach to discounting has attractive properties, especially in the face of the intergenerational challenge of climate change; for these reasons and others, RDGU may prove the best family of social orderings to choose. Or perhaps, following Blackorby, Bossert, and Donaldson's (1995) recognition of independence of utilities of the dead, we may decide that separability or existence independence makes the CLGU family best—without necessarily deciding whether the critical level is zero or positive.

A final, quantitatively-minded alternative is to emphasize the large number of policy questions where qualitatively-distinct social orderings would agree. With RDGU's β close to 1 and CLGU's critical level close to zero, the policy recommendations of these two approaches and total utilitarianism will agree in practice. Indeed, policy evaluations routinely investigate the robustness of conclusions to a range of functional forms and normative parameters, such as time discounting, inequality aversion, or values of a statistical life. Ultimately, population economics can similarly verify the robustness of policy conclusions to alternative shapes of f, g, and h or values of β and c and to alternative social orderings—each of which would imply "repugnant" conclusions in some imaginable case.

Appendix

Proofs

Proof of Theorem 1

The proof uses Aggregation twice. If $\succsim$ satisfies Axiom 5, then choose $v \in (0, \varepsilon)$; this will be used to construct a $\mathbf{v}$ without negative lives. If $\succsim$ satisfies Axiom 6 but not Axiom 5, choose $v < 0$, which implies $v < \varepsilon$. Choose $\delta \in (0, \min(\frac{\varepsilon - v}{3}, \frac{\varepsilon}{3}))$ and, if $v < 0$, $\delta < -v$. For any m, construct $\left(u^h \mathbf{1}_{n^h}, v\mathbf{1}_m\right)$; if m is large enough then $(v + \delta)\mathbf{1}_{n^h+m} \succ \left(u^h \mathbf{1}_{n^h}, v\mathbf{1}_m\right)$, by Aggregation. For any n^ε, construct $\left(\varepsilon\mathbf{1}_{n^\varepsilon}, u^\ell \mathbf{1}_{n^\ell}, v\mathbf{1}_m\right)$; if n^ε is large enough (once m is fixed) then $\left(\varepsilon\mathbf{1}_{n^\varepsilon}, u^\ell \mathbf{1}_{n^\ell}, v\mathbf{1}_m\right) \succ (\varepsilon - \delta)\mathbf{1}_{n^\varepsilon+n^\ell+m}$. Choose m large enough to satisfy Aggregation; then choose n^ε large enough to satisfy Aggregation and such that $n^\varepsilon + n^\ell + m > n^h + m$. By one of the sign axioms and by the construction of δ, $(\varepsilon - \delta)\mathbf{1}_{n^\varepsilon+n^\ell+m} \succ (v + \delta)\mathbf{1}_{n^h+m}$. Then by applying transitivity twice, $\left(\varepsilon\mathbf{1}_{n^\varepsilon}, u^\ell \mathbf{1}_{n^\ell}, v\mathbf{1}_m\right) \succ \left(u^h \mathbf{1}_{n^h}, v\mathbf{1}_m\right)$.

Footnote 21 (continued)

this Wide Collective Principle. This claim made two mistakes. We cannot justifiably reject strong arguments merely by claiming that their conclusions are implausible" (p. 154).

Proof of Proposition 1

Because the extended very repugnant conclusion implies the very repugnant conclusion, the proof of Theorem 1 applies for orderings that satisfy Aggregation and a sign axiom.

For social orderings that satisfy Aggregation but not a zero axiom, or for social orderings that merely satisfy Minimal Equality Preference, the proof by construction uses the $u_j = v_j + \varepsilon$ horn of the definition of ε-change. For Aggregation-satisfying orderings, simply include a very large base population; improve many lives by ε.

To begin a construction for Non-Aggregation social orderings, choose any ε, u^ℓ, u^h, n^ℓ, and n^h according to the Extended VRC. Next, set m^h and m^ℓ in the EVRC to both be the maximum of $n^h + 1$ and $n^\ell + 1$. Then, let ξ in the Axiom be $u^\ell - \varepsilon$ from the EVRC. Let δ in the axiom be ε from the EVRC. Notice that $\xi + \delta$ in the Axiom now equals u^ℓ from the EVRC. Now, let $\mathbf{u}$ from the Axiom be $u^h \mathbf{1}_{m^h}$. Notice that $n(\mathbf{u})$ from the Axiom is now fixed at $n(\mathbf{u}) = m^h = m^\ell$ from the EVRC.

The construction next uses the Minimal Equality Preference axiom. We have now specified a $\mathbf{u}$, ξ, and δ. So there exists an n^* such that if $n > n^*$ then $(\xi + \delta)\mathbf{1}_{n+n(\mathbf{u})} \succ \left(\xi \mathbf{1}_n, \mathbf{u}\right)$. Choose such an n and call it $\tilde{n}$. This construction fulfills the conditions of the Extended VRC. Note that we may choose any $\mathbf{v} \in \Omega$. Let $\mathbf{v} = \xi \mathbf{1}_{\tilde{n}}$. Now notice that $\left(\xi \mathbf{1}_{\tilde{n}}, \mathbf{u}\right)$ from the Axiom is $\left(\xi \mathbf{1}_{\tilde{n}}, u^h \mathbf{1}_{m^h}\right)$, which is $\left(\mathbf{v}, u^h \mathbf{1}_{m^h}\right) = \mathbf{v}^h$ from the EVRC. Let $\mathbf{v}^\ell = (\xi + \delta)\mathbf{1}_{\tilde{n}+n(\mathbf{u})} = (\xi + \delta)\mathbf{1}_{\tilde{n}+m^\ell} = \left(u^\ell \mathbf{1}_{\tilde{n}}, u^\ell \mathbf{1}_{m^\ell}\right) = \left((\xi + \varepsilon)\mathbf{1}_{\tilde{n}}, u^\ell \mathbf{1}_{m^\ell}\right)$. Set n^ε equal to $\tilde{n}$. Finally, we can see that $\left((\xi + \varepsilon)\mathbf{1}_{\tilde{n}}, u^\ell \mathbf{1}_{m^\ell}\right)$ is separated by n^ε ε-changes from $\left(\mathbf{v}, u^\ell \mathbf{1}_{m^\ell}\right)$.

Maximin and maximax both can be shown to imply the extended very repugnant conclusion by having $\mathbf{v}$ contain the least or greatest (respectively) utility level, and then increasing this with one ε-change.

References

Adler MD (2009) Future generations: a prioritarian view. George Wash Law Rev 77(5/6):1478

Anglin B (1977) The repugnant conclusion. Can J Philos 7(4):745–754

Arrhenius G (n.d.) Population ethics: the challenge of future generations. (unpublished typescript, July 2017)

Arrhenius G (2000) An impossibility theorem for welfarist axiologies. Econ Philos 16(2):247–266

Arrhenius G (2003) The very repugnant conclusion. Uppsala Philosophical Studies, p 51

Arrhenius G, Ryberg J, Tännsjö T (2017) The repugnant conclusion. In: Zalta EN (ed) The Stanford encyclopedia of philosophy. Stanford University, Spring 2017 edition

Asheim GB, Zuber S (2014) Escaping the repugnant conclusion: rank-discounted utilitarianism with variable population. Theor Econ 9(3):629–650

Asheim GB, Zuber S (2018) Rank-discounting as a resolution to a dilemma in population ethics. Working paper, University of Oslo and PSE

Blackorby C, Donaldson D (1984) Social criteria for evaluating population change. J Public Econ 25(1–2):13–33

Blackorby C, Bossert W, Donaldson D (1995) Intertemporal population ethics: critical-level utilitarian principles. Econometrica 63(6):1303–1320

Blackorby C, Bossert W, Donaldson D, Fleurbaey M (1998a) Critical levels and the (reverse) repugnant conclusion. J Econ 67(1):1–15

Blackorby C, Primont D, Russell RR (1998b) Separability: a survey. Handb Util Theory 1:49–92
Blackorby C, Bossert W, Donaldson DJ (2005) Population issues in social choice theory, welfare economics, and ethics. Cambridge University Press
Bossert W (2017) Anonymous welfarism, critical-level principles, and the repugnant and sadistic conclusions. Working paper, University of Montréal
Boucekkine R, Fabbri G (2013) Assessing Parfit's repugnant conclusion within a canonical endogenous growth set-up. J Popul Econ 26(2):751–767
Broome J (2004) Weighing lives. Oxford UP, Oxford
Broome J (2012) Climate matters: ethics in a warming world (Norton global ethics series). WW Norton & Company
Broome J (2015) Equality versus priority: a useful distinction. Econ Philos 31(2):219–228
Budolfson M, Spears D (2018) Why the repugnant conclusion is inescapable. Working paper, Princeton University CFI
Carlson E (forthcoming) On some impossibility theorems in population ethics. In: Campbell T et al (ed) The Oxford handbook of population ethics. Oxford
Cowen T (1996) What do we learn from the repugnant conclusion? Ethics 106(4):754–775
Cowen T (2018) Stubborn attachments. Stripe Press
Dasgupta P (2005) Regarding optimum population. J Polit Philos 13(4):414–442
Dasgupta P (2016) Birth and death. Working paper 2016/19, Cambridge-INET
Deschamps R, Gevers L (1977) Separability, risk-bearing, and social welfare judgements. Eur Econ Rev 10(1):77–94
Fleurbaey M (2010) Assessing risky social situations. J Polit Econ 118(4):649–680
Fleurbaey M, Tungodden B (2010) The tyranny of non-aggregation versus the tyranny of aggregation in social choices: a real dilemma. Econ Theory 44(3):399–414
Fleurbaey M, Zuber S (2015) Discounting, beyond utilitarianism. Economics 9(2015–12):1–52
Franz N, Spears D (2020) Mere addition is equivalent to avoiding the sadistic conclusion in all plausible variable-population social orderings. Econ Lett 196:109547
Gustafsson JE (2017) Moral aggregation, by Iwao Hirose. Mind 126(503):964–967
Hurka T (1983) Value and population size. Ethics 93(3):496–507
Lawson N, Spears D (2018) Optimal population and exhaustible resource constraints. J Popul Econ 31(1):295–335
McCarthy D, Mikkola K, Thomas T (2020) Utilitarianism with and without expected utility. J Math Econ. 87:77–113. https://doi.org/10.1016/j.jmateco.2020.01.001
Nerlove M, Razin A, Sadka E (1982) Population size and the social welfare functions of Bentham and Mill. Econ Lett 10(1–2):61–64
Ng Y-K (1989) What should we do about future generations?: impossibility of Parfit's theory X. Econ Philos 5(2):235–253
Palivos T, Yip CK (1993) Optimal population size and endogenous growth. Econ Lett 41(1):107–110
Parfit D (1984) Reasons and persons. Oxford UP, Oxford
Parfit D (2016) Can we avoid the repugnant conclusion? Theoria 82(2):110–127
Parfit D (2017) Future people, the non-identity problem, and person-affecting principles. Philos Public Aff 45(2):118–157
Pivato M (2020) Rank-additive population ethics. Econ Theory 69:861–918
Roberts MA (2011) The asymmetry: a solution. Theoria 77:333–367
Scovronick N, Budolfson MB, Dennig F, Fleurbaey M, Siebert A, Socolow RH, Spears D, Wagner F (2017) Impact of population growth and population ethics on climate change mitigation policy. Proc Natl Acad Sci 114(46):12338–12343
Sider TR (1991) Might theory X be a theory of diminishing marginal value? Analysis 51(4):265–271
Spears D (2017) Making people happy or making happy people? Questionnaire-experimental studies of population ethics and policy. Soc Choice Welf 1–25
Temkin LS (2014) Rethinking the good: moral ideals and the nature of practical reasoning. Oxford University Press
Zuber S, Asheim GB (2012) Justifying social discounting: the rank-discounted utilitarian approach. J Econ Theory 147(4):1572–1601
Zuber S, Venkatesh N, Tännsjö T, Tarsney C, Orri SH, Steele K, Spears D, Sebo J, Pivato M, Ord T, Ng Y-K, Masny M, MacAskill W, Lawson N, Kuruc K, Hutchinson M, Gustafsson JE, Greaves H, Forsberg L, Fleurbaey M, Coffey D, Cato S, Castro C, Campbell T, Budolfson M, Broome J, Berger

A, Beckstead N, Asheim GB (2021) What should we agree on about the repugnant conclusion?. Utilitas. https://doi.org/10.1017/S095382082100011Xgg

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Dean Spears[1,2,3,4,5] · **Mark Budolfson**[6,7,8]

1 Economics Department and Population Research Center, University of Texas at Austin, Austin, USA

2 Economics and Planning Unit, Delhi Centre, Indian Statistical Institute, New Delhi, India

3 IZA, Bonn, Germany

4 Institute for Futures Studies, Stockholm, Sweden

5 r.i.c.e, Lucknow, India

6 Department of Environmental and Occupational Health and Justice, Rutgers University, New Brunswick, USA

7 Center for Population-Level Bioethics, Rutgers University, New Brunswick, USA

8 Department of Philosophy, Rutgers University, New Brunswick, USA