I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 12668

# Why Variable-Population Social Orderings Cannot Escape the Repugnant Conclusion: Proofs and Implications

Dean Spears
Mark Budolfson

OCTOBER 2019

DISCUSSION PAPER SERIES

IZA DP No. 12668

# Why Variable-Population Social Orderings Cannot Escape the Repugnant Conclusion: Proofs and Implications

**Dean Spears**
*University of of Texas at Austin, Economics and Planning Unit, Indian Statistical Institute - Delhi Centre, IZA, Institute for Future Studies and r.i.c.e.*

**Mark Budolfson**
*UVM and ANU*

OCTOBER 2019



ISSN: 2365-9793

**IZA – Institute of Labor Economics**

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

# ABSTRACT

## Why Variable-Population Social Orderings Cannot Escape the Repugnant Conclusion: Proofs and Implications

The population literature in theoretical economics has long focused on attempts to avoid the repugnant conclusion. We advance the literature by proving that no social ordering in population economics can escape the repugnant conclusion in all instances. As we show, prior results depend on a formal definition of the repugnant conclusion that artificially excludes some repugnant cases. In particular, the literature traditionally formalizes the repugnant conclusion to exclude cases that include an unaffected subpopulation. We relax this normatively irrelevant exclusion, and others. We prove that any candidate social ordering that satisfies either a basic axiom of Aggregation or Non-Aggregation implies some instance of the repugnant conclusion. Therefore, the repugnant conclusion provides no methodological guidance for theory or policymaking, because it cannot discriminate among candidate social orderings. This result is of practical importance because evaluation of important climate or development policies depends on comparing social welfare across populations of differing sizes.

**JEL Classification:** J10, J13, J18, D63

**Keywords:** population economics, population ethics, repugnant conclusion, RDGU, CLGU, separability

**Corresponding author:**
Dean Spears
Department of Economics
The University of Texas at Austin
2225 Speedway
BRB 1.116, C3100
Austin, Texas 78712
USA

E-mail: dspears@utexas.edu

# 1 Introduction: Which conclusions are repugnant?

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 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 (Blackorby and Donaldson, 1984; Dasgupta, 1995).

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." 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 sums wellbeing: $V^{TU}(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} u_i$). Average utilitarianism, for example, (where $V^{AU}(\mathbf{u}) = \frac{1}{n(\mathbf{u})} \sum_{i=1}^{n(\mathbf{u})} u_i$) is cannonically interpreted not to imply the repugnant conclusion.

Our study clarifies the implications of the repugnant conclusion. The leading papers in population economics have focused on proving formal theorems about which families of social welfare functions do or do not imply the repugnant conclusion (Ng, 1989; 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 at a large theoretical cost.[1]

This paper contributes theorems and examples that reveal that this understanding of

[1]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 (n.d.) and Bossert (2017).

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, it is an implication of *every* social ordering in population economics.

To reach this conclusion, we show that the literature, in a history of path-dependence from prominent original examples, has used several formal definitions of the repugnant conclusion, all of which capture only a subset of equivalently repugnant cases. We build on Fleurbaey and Tungodden (2010), who demonstrate a dilemma in welfare economics between fundamentally *aggregative* social welfare functions (such as utilitarianism) which consider every person's interests and fundamentally *non-aggregative* social welfare functions (such as maximin) which attend most or only to the worst-off. We show that whether a social welfare function fulfills an axiom of Aggregation or Non-Aggregation, it must imply a single, extended very repugnant conclusion which preserves all of the repugnance of standard formalizations of the repugnant conclusion in the literature.

The implication is that population ethics is not fundamentally a choice between the repugnant conclusion or other undesirable implications. This only appears to be the case because of a history of formal definitions that exclude many instances of repugnance. Similarly, repugnance is not a defining property of total utilitarianism: for example, there are choice sets from which average utilitarianism makes "repugnant" choices of worse lives rather than better lives, but total utilitarianism does not.[2] We conclude that because repugnance cannot be escaped, escaping it should not be a goal, so there is no normative reason to impose an axiomatic requirement for population ethics to avoid the repugnant conclusion.

## 1.1 Our contribution, in the context of prior impossibility results

What is essential to the repugnance of the repugnant conclusion? What separates mere counterexamples against a social ordering from repugnant conclusions? To motivate our paper, we offer an introductory illustration of the efforts to escape Parfit's version of the repugnant conclusion. A recent important advancement of this literature is Asheim and Zuber's (2014) Rank-Dependent Generalized Utilitarianism (RDGU), which we use for illustration throughout this paper. This social welfare function transforms each person's utilities by an increasing function $g$ (which may be linear or concave), and weights utilities

[2] 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).

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 $\beta \in (0,1)$ and the square-bracket index $[r]$ indicates that $\mathbf{u}$ is ordered in increasing rank. As Asheim and Zuber prove, RDGU escapes Parfit's formulation of the repugnant conclusion.[3] Although RDGU escapes Parfit's original example of repugnance, it has the following unintuitive implication, for a comparison of same-sized populations:

**Example 1.** *Let $\varepsilon > 0$ be any small utility increment, $r^* > 1$ be any counting rank, $m > r^*$ be any large number of people, and $\delta > 0$ (and $\delta > \varepsilon$) be any large utility increment. Then, there exists same-sized populations $\mathbf{u}$ and $\mathbf{v}$ such that:*

- *Only rank $r^*$ is strictly better-off under $\mathbf{v}$ than $\mathbf{u}$, and only by $\varepsilon$;*
- *For all other ranks $r \neq r^*$, including the worst-off person, $\mathbf{v}$ is strictly worse than $\mathbf{u}$; and*
- *For at least $m$ people ($m$ rank orders $r$), $\mathbf{v}$ is much worse: $v_{[r]} + \delta \leq u_{[r]}$,*

*such that RDGU strictly prefers $\mathbf{v}$ to $\mathbf{u}$, but $\mathbf{u}$ is strictly preferred by every population-sensitive social welfare function in the population economics literature, incuding maximin, maximax, total and average utilitarianism and prioritarianism, and critical-level generalized utilitarianism.*

We see this as an unintiutive implication of RDGU. But Example 1 may plausibly be argued merely to be a counterexample to RDGU that is separate and distinct from *the repugnant conclusion*, because Example 1 does not concern quantity-quality tradeoffs, which are the focus of the repugnant conclusion. But consider Example 2, which does make a quantity-quality tradeoff, and in particular makes the social choice of lower-quality lives over higher-quality lives:

**Example 2.** *Let $\varepsilon > 0$ be any small utility increment, $u^h > 0$ be any very high utility level, and $n^h > 0$ be any large number of lives at $u^h$. Then, there exists a number of $\varepsilon$-quality lives $n^\varepsilon > 0$ and an unaffected, intersecting sub-population $\mathbf{v}$ such that adding $n^\varepsilon$ lives each at $\varepsilon$ to $\mathbf{v}$ is preferred by RDGU to adding $n^h$ excellent lives at $u^h$.*

Both examples are proven in appendix section A.1. Example 2 shows that RDGU implies repugnance, after all. RDGU has this implication because, as we show, every

[3]In the context of Ng's (1989) theorem, RDGU does this by denying the Mere Addition axiom.

plausible social welfare function *must*. There is no *normative* difference between, on the one hand, any repugnance of a social ordering recommending the creation of worse lives instead of better lives in Example 2 and, on the other hand, any repugnance in Parfit's example of the repugnant conclusion.

What does Example 2 tell us? The literature has long recognized the impossibility of a population-sensitive social welfare function that fully accords with every desirable intuition (*e.g.*Ng, 1989; Arrhenius, 2000; Blackorby et al., 2005). But this literature would not consider Example 2 to be an instance of the repugnant conclusion: the presence of the unaffected sub-population v causes Example 2 to be excluded from standard formal definitions of the repugnant conclusion in the literature. However, *every* policy choice includes unaffected sub-populations, including, at a minimum, the set of people who have already died (Blackorby et al., 1995). As we show below, the literature has also excluded other instances of repugnance that differ from traditional formalizations only in other normatively unimportant ways. Only because these examples are excluded from the habitual formalization of repugnance are some social orderings (such as RDGU) understood to escape the repugnant conclusion. Because these equivalently-repugnant cases should not be excluded, the repugnant conclusion cannot be escaped. Escaping it, therefore, should not be a theoretical goal — because it is impossible.

### 1.2 Outline

Section 2 introduces our setting and four basic axioms. These have been shown in the literature to be sufficient for a reduced-form representation of a social ordering as an equally-distributed equivalent level of well-being and population size. Following the population economics literature, we only consider the very large set of social orderings with such a reduced-form representation.

Section 3 introduces a fundamental question of this paper: what is repugnance? What understandings of the repugnant conclusion fully capture the repugnance of Parfit's example? How should such repugnance be translated into a family of formal definitions? We observe that, although the prior literature includes several "repugnant conclusions," 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. Here, we define *unrestricted* versions of the repugnant conclusion and the very repugnant conclusion, which may or may not include such unaffected lives. 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 no special problem for population ethics.

We next consider two large subsets of social orderings, which include all that we know to be advanced in the population economics literature and more. Section 4 introduces an Aggregation axiom. All social orderings which satisfy the Aggregation axiom (and a definition for zero) imply the very repugnant conclusion, in our unrestricted version. Because Aggregation is closely related to separability in same-number, risk-free cases, this includes many social orderings — among them many that are commonly understood to escape repugnance.

Section 5 introduces a Non-Aggregation axiom. This axiom reflects orderings such as maximin or RDGU which give unequal emphasis to a subset of lives. We show that all social orderings with a reduced-form representation which satisfy either Aggregation or Non-aggregaation (with no other requirements) imply the extended very repugnant conclusion. Either in aggregating large quantities of tiny changes, or in ignoring them, however many, 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 and basic axioms

We largely use the same notation for welfarist, variable-population social evaluation used by Blackorby et al. (2005). $\mathbb{Z}$ are the integers, $\mathbb{R}$ are 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, where the $i$th position in the vector $u_i$ is the lifetime utility of person $i$.[4] Following Asheim and Zuber (2014), when an index is inclosed in square brackets it 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. The size of $\mathbf{u}$

[4] We only consider what Pivato (2018) calls *actualist* populations, considering only the utility levels of people who come to exist; Pivato contrasts those with *possibilist* populations, which assign zero utility to possible people who never exist.

is $n(\mathbf{u}) \in \mathbb{Z}_{++}$, so $u \in \mathbb{R}^{n(\mathbf{u})}$. In comparing the utilities in populations, $\mathbf{u} \geq \mathbf{v}$ means that $u_i \geq v_i$ for all $i$; $\mathbf{u} > \mathbf{v}$ means that $u \geq v$ 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 unit vector, so $\xi \mathbf{1}_n$ is a population in which all $n$ people have equal utility $\xi$. Applied to individual populations, $\cup$ notation combines populations, so $n(\mathbf{u} \cup \mathbf{v}) = n(\mathbf{u}) + n(\mathbf{v})$.

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 $\Omega$. $\succsim$ is a binary relation 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, unbounded, and linear or 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. Not every case makes nontrivial use of $f$ or $g$, so these are omitted for clarity when they are identify functions.

We ignore risk and uncertainty, focusing only on social orderings of degenerate outcomes. Some questions in social welfare that are debated in risky cases are less controversial in risk-free cases, such as same-number, risk-free separability (Fleurbaey, 2010; Broome, 2015).

We begin with a set of basic axioms on $\succsim$. Although the philosophical literature on population ethics contains papers which explore denying each of these axioms, these have been uncontroversial in the economics literature since Blackorby and Donaldson (1984), and we adopt them throughout the paper here.

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

**Axiom 2** (Anonymity). *For all $\mathbf{u}, \mathbf{v} \in \Omega$ such that $n(\mathbf{u}) = n(\mathbf{v})$, if there exists a bijection $\rho : \{1, \ldots, n(\mathbf{u})\} \to \{1, \ldots, n(\mathbf{u})\}$ 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 summarized 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 social welfare function $V : \Omega \to \mathbb{R}$ and a reduced-form social welfare funciton $W : \mathbb{Z}_{++} \times \mathbb{R} \to \mathbb{R}$, such that for all $\mathbf{u}, \mathbf{v} \in \Omega$,*

$$\mathbf{u} \succsim \mathbf{v} \equiv V(\mathbf{u}) = W\left(n(\mathbf{u}), \Xi(\mathbf{u})\right) \geq W\left(n(\mathbf{v}), \Xi(\mathbf{v})\right) = V(\mathbf{v}),$$

*where the restriction of $\mathbf{v}$ to $\mathbb{R}^n$ is continuous for all $n \in \mathbb{Z}_{++}$, $W$ is continuous and increasing in its second argument, and the EDE $\Xi$ 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 $\mathbf{u}$.*

Every social ordering that has received attention in the population economics literature [5] can be expressed in this reduced form:

$$\begin{array}{rclrclrcl} W^{AU} & = & \Xi. & W^{TU} & = & n\Xi. & W^{TPri} & = & ng(\Xi). \\ W^{X'} & = & f(n)\Xi. & W^{CLGU} & = & n\left(g(\Xi) - g(\alpha)\right). & W^{RDGU} & = & \frac{\beta(1-\beta^n)}{1-\beta} g(\Xi). \end{array}$$

CLGU is critical-level generalized utilitarianism (Blackorby and Donaldson, 1984), $X'$ is variable-value utilitarianism (Ng, 1989), and TPri stands for total prioritarianism (Adler, 2009). Maximin does not satisfy axiom 4 (Pareto), but can still be reduced as $W^M = \Xi$.

# 3 Repugnant conclusions

Whether the repugnant conclusion can be avoided depends on what the repugnant conclusion is. Because Parfit's (1984) original statement of the repugnant conclusion specifically invokes a well-off population of ten billion people, every paper in the formal population literature has adopted a technical definition that goes beyond Parfit's original example. These definitions sometimes disagree, and even Parfit (2016) has written about the possible heterogeneity in instances of repugnant conclusions.[6] Yet, the population economics

[5] In the philosophical literature, for example, Temkin (2014) considers denial of the transitive part of Axiom 1, Roberts (2011) denies Axiom 2, and Carlson (2017) denies Axiom 3. Such issues are a focus of a companion working paper in that literature Budolfson and Spears (2018), which does not contain the formal results of this paper.

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

literature typically formalizes the repugnant conclusion as:[7]

**Definition 1** (The (original, restricted) repugnant conclusion)**.** For any large $u^h \in \mathbb{R}_{++}$, any large $n^h \in \mathbb{Z}_{++}$, and any small, positive $\varepsilon > 0$, there exists $n^\varepsilon \in \mathbb{Z}_{++}$ such that $\varepsilon \mathbf{1}_{n^\varepsilon} \succ u^h \mathbf{1}_{n^h}$ (Parfit, 1984).

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.'s (1995) "independence of the utilities of the dead" axiom is that populations exist in time, and past populations cannot be influenced by future choices.[8] The existence of a dead, past sub-population — or of any other unaffected population — is irrelevant to any repugnance in the choice to create $n^\varepsilon$ lives at $\varepsilon$ 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 large $u^h \in \mathbb{R}_{++}$, any large $n^h \in \mathbb{Z}_{++}$, and any small, positive $\varepsilon > 0$, there exists $n^\varepsilon \in \mathbb{Z}_{++}$ and $\mathbf{v} \in \Omega \cup \{\emptyset\}$ such that $\varepsilon \mathbf{1}_{n^\varepsilon} \cup \mathbf{v} \succ u^h \mathbf{1}_{n^h} \cup \mathbf{v}$.

Thus, our contribution begins by offering a revised formalization of the repugnant conclusion. In our terminology, definition 2 is *the* repugnant conclusion.[9] Definition 1 is the *restricted* repugnant conclusion.

[7]In fact, in their footnote 2, Asheim and Zuber (2014) note that another formal definition of the repugnant conclusion is available which would be implied by a slightly different set of social orderings. Therefore, even the prior literature has noted ambiguity in what the repugnant conclusion *is* (Dasgupta, 2005). Further substantive variation in the formalization of the repugnant conclusion in the literature is discussed in Section 4.4.

[8]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 we find this axiom plausible, it is unrelated to our argument in this paper.

[9]Arrhenius (n.d.) uses an equivalent condition called the Strong Quality Addition Principle; Anglin (1977)

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 |

Arrhenius (2003) introduced the *very* repugnant conclusion, which intensified the repugnant conclusion by stipulating that the $\varepsilon$ 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 large $u^h \in \mathbb{R}_{++}$, any large $n^h \in \mathbb{Z}_{++}$, any very negative $u^\ell \in \mathbb{R}_{--}$, any large $n^\ell \in \mathbb{Z}_{++}$, and any small, positive $\varepsilon > 0$, there exists $n^\varepsilon \in \mathbb{Z}_{++}$ and $\mathbf{v} \in \Omega \cup \{\emptyset\}$ such that $\varepsilon \mathbf{1}_{n^\varepsilon} \cup u^\ell \mathbf{1}_{n^\ell} \cup \mathbf{v} \succ u^h \mathbf{1}_{n^h} \cup \mathbf{v}$.

Although we consider Definition 3 to be a natural formalization of any repugnance in quality-quantity tradeoffs, 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 with Table 1, which summarizes this paper's revision and generalization of terminology in the literature for repugnant conclusions.

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. Moreover, 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

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 theorems below use a different condition, the unrestricted very repugnant conclusion.

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

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 $n(\mathbf{v}) = 0$, and that there is no normative or practical reason to impose a constraint to the restricted subset.

# 4 Consequences of Aggregation

This paper partitions the social orderings defended in the population economics literature into those that satisfy an axiom of Aggregation and those that satisfy an axiom of non-Aggregation. This section focuses on the former; the next section focuses on the latter. Here, section 4.1 introduces axioms which make meaningful the zero level of utility (used in the repugnant conclusion's emphasis on $\varepsilon > 0$ lives). Then, section 4.2 presents the Aggregation axiom and, with it, our first theorem. Finally, section 4.3 discusses properties that are sufficient for a social ordering to satisfy Aggregation.

## 4.1 Axioms of zero

The repugnant conclusion invokes the prospect of lives at $\varepsilon > 0$, slightly-positive "lives that are barely worth living" in Parfit's words. 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."

The classic zero axiom, named "mere addition" by Parfit, is that adding a life of utility above zero 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 $\mathbf{v} \cup u\mathbf{1}_1 \succsim \mathbf{v}$.

Many social orderings in the literature do not satisfy mere addition. Average utilitarianism and Ng's (1989) variable-value utilitarianism fail mere addition because additional positive lives could lower average utility. Here, we focus again on RDGU as a leading recent proposal that avoid's Parfit's restricted example of the repugnant conclusion. RDGU not only fails mere addition, but additionally may refuse the addition of arbitrarily good,

problem we face. We are here."

positive lives even when such additional lives increase average utility (or $g$-transformed utility):

**Example 3.** *Let $n^h \in \mathbb{Z}_{++}$ be any number of lives and $u^h \in \mathbb{R}_{++}$ be any large utility level. Then there exists a $\mathbf{v} \in \Omega$ such that $\mathbf{v} \gg 0$ and $\mathbf{v} \cup u^h \mathbf{1}_{n^h}$ has higher total and average utility than $\mathbf{v}$ (and the same minimum), but RDGU ranks $\mathbf{v} \succ \mathbf{v} \cup u^h \mathbf{1}_{n^h}$.*

Note that Example 3 is not true of maximin, maximax, total utilitarianism, average utilitarianism, Ng's variable-value utilitarianism, or CLGU. In fact, RDGU also implies the following consequence, which is arguably more striking, and which violates what Arrhenius (n.d.) calls a "dominance" condition, in which good lives are added while improving the welfare of all otherwise-existing people:

**Example 4.** *Let $n^h \in \mathbb{Z}_{++}$ be any number of lives and $u^h \in \mathbb{R}_{++}$ be any large utility level. Then there exists a $\mathbf{v} \in \Omega$ such that $\mathbf{v} \gg 0$ and $\mathbf{v} \cup u^h \mathbf{1}_{n^h}$ has higher total and average utility than $\mathbf{v}$ (and the same minimum) and there further exists $\varepsilon \in \mathbb{R}_{++}$ such that RDGU ranks $\mathbf{v} \succ \mathbf{v}' \cup u^h \mathbf{1}_{n^h}$, where $v'_i = v_i + \varepsilon$ for all $i$.*

Examples 3 and 4 are proven in appendix A.2. Therefore, instead of using mere addition, our Theorem 1 allows a social ordering to satisfy either of two even more attractive zero axioms. Average utilitarianism, variable-value utilitarianism, and RDGU each satisfy axioms 5 and 6 (as do total utilitarianism and prioritarianism; as do orderings such as maximin and maximax with reduced forms that are insensitive to $n$). So, axioms 5 and 6 can be satisfied even by social orderings with the implications in Examples 2 and 3. Axioms 5 and 6 use the equally-distributed equivalent to avoid the problematic cases for average-type theories where additional positive lives bring down the population-wide average.

**Axiom 5** (First zero axiom: EDE dominance)**.** *For any $\mathbf{u}, \mathbf{v} \in \Omega$ with a reduced-form representation, if $\Xi(\mathbf{u}) > \Xi(\mathbf{v}) > 0$ and $n(\mathbf{u}) > n(\mathbf{v})$, then $\mathbf{u} \succ \mathbf{v}$.*

**Axiom 6** (Second zero axiom: EDE priority for lives worth living)**.** *For any $u, v \in \Omega$ with a reduced-form representation, if $\Xi(\mathbf{u}) > 0 > \Xi(\mathbf{v})$, then $\mathbf{u} \succ \mathbf{v}$.*

If the reader finds axioms expressed in terms of the EDE difficult to assess intuitively, note that, because the social orderings can be represented with a reduced form, it would be equivalent to express axioms 5 and 6 limited only to the $\mathbb{R} \times \mathbb{Z}_{++}$ subset of $\Omega$ containing only perfectly equal-utility populations.

## 4.2 Axiom of Aggregation

The last step before Theorem 1 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) (who focus on a limited-size loss, rather than a loss accruing to a limited fraction of the population), 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 small positive real number $\delta > 0$, and any utility level $\xi \in \mathbb{R}$, there exists $n^* \in \mathbb{Z}_{++}$ such that if $n > n^*$, then $\left|\Xi\left(\mathbf{u} \cup \xi \mathbf{1}_n\right) - \Xi\left(\xi \mathbf{1}_{n+n(\mathbf{u})}\right)\right| < \delta$.*

The Aggregation axiom holds that the EDE 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 between $\mathbf{u} \cup \xi \mathbf{1}_n$ and $\xi \mathbf{1}_{n+n(\mathbf{u})}$ holds population size constant. As section 4.3 details, many social orderings in the population economics literature satisfy Aggregation.

**Theorem 1.** *If $\succsim$ satisfies the basic axioms (1-4) or otherwise has a reduced-form representation with the properties in the Lemma, Aggregation (7), and at least one of the zero 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 $\mathbf{v} \gg 0$).*

*Proof.* See appendix section A.3. □

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, TU would make a repugnant choice are not a superset of the choices in which AU would make a repugnant choice.[11] Nor, as the theorem notes, is it required that $\mathbf{v} \ll 0$.

## 4.3 Which social welfare functions satisfy Aggregation?

The significance of Theorem 1 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

[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 $n^\ell = 0$). AU would choose the $\varepsilon$ lives and TU would choose the $u^h$ lives. If $\varepsilon = 0$, which would only increase the repugnance of that choice, AU would continue to choose the $\varepsilon$ lives as $n^\varepsilon$ becomes ever larger, but TU would continue to choose the $u^h$ lives.

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).[12] 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, which 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. Theorem 1 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 form:

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

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

**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* $\mathbf{u} \cup \mathbf{w} \succsim \mathbf{v} \cup \mathbf{w}$ *then* $\mathbf{u} \cup \mathbf{x} \succsim \mathbf{v} \cup \mathbf{x}$ *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. This is true, for example, of RDGU. Imagine a policy that would take 10

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

[13] 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 Theorem 1 and therefore imply the very repugnant conclusion. Nothing hinges on our use of this terminology from the literature, however.

units of utility away from each of 100 people at utility 100 and give 8 units of utility to each of 100 people at utility 10. Is this policy desirable? Any same-number additively-separable ordering can answer immediately, but RDGU additionally needs to know how many rank-orders (perhaps people in different continents, or ancient Egypt, or the far future) are between utilities 10 and 100. We find this both normatively and pragmatically implausible.

Same-number independence is not quite sufficient for the functional form in equation 1. As Blackorby et al. (1998) 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 $\Xi$ 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:

**Axiom 9** (Replication invariance)**.** *For any* $\mathbf{u}, \mathbf{v} \in \Omega$ *and any* $n \in \mathbb{Z}_{++}$*, if* $\mathbf{u} \succsim \mathbf{v}$ *then* $\bigcup_n \mathbf{u} \succsim \bigcup_n \mathbf{v}$ *and similarly for* $\succ$*.*

Note that replication invariance does not itself imply same-number separability (maximin satisfies replication invariance but not separability) nor the reverse. Although many of the social orderings in population economics satisfy replication invariance, RDGU does not.[14] Fleurbaey and Tungodden (2010) and Asheim and Zuber (2018) emphasize the role of replication invariance in the conflict between aggregation and non-aggregation. In our framework, 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 $u \in \mathbb{R}$. Then

$$\lim_{m\to\infty}\left(\frac{1}{m+n(\mathbf{u})}\sum_{i=1}^{n(\mathbf{u})} g^{m+n(\mathbf{u})}(u_i) + \frac{m}{m+n(\mathbf{u})} g^{m+n(\mathbf{u})}(\xi)\right) - g^{m+n(\mathbf{u})}(\xi) = 0, \tag{3}$$

for all $\mathbf{u}$ and $\xi$, which is what Aggregation requires. So, for example, setting $g^n(u_i) = u_i\sqrt{n} - e^{-u_i} + 1$ in equation 2 would satisfy Aggregation and every other condition in

[14] With $\beta = 0.7$, RDGU ranks $(-1, 1.5) \succ 0\mathbf{1}_2$ but $(-1, -1, 1.5, 1.5) \prec 0\mathbf{1}_4$.

Theorem 1, and therefore would imply the very repugnant conclusion, but would not satisfy replication invariance.[15] We are aware of no same-number separable social welfare function that is defended in the literature that does not satisfy equation 3, nor of any normative argument for violating it in a same-number separable social ordering.

**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); if additionally it satisfies at least one of the zero 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 in equation 3 that same-number inequality aversion (or, more broadly, how the increasing shape of $g^n(u)$ changes in $n$) is not changing too quickly in population size. A very broad family of social welfare functions is:

$$V(\mathbf{u}) = f\left(n\left(\mathbf{u}\right)\right)\left[h\left(\frac{1}{n(\mathbf{u})}\sum_{i=1}^{n(\mathbf{u})} g(u_i)\right) - h\left(g(c)\right)\right], \tag{4}$$

where, in addition to the definitions elsewhere, $c \in \mathbb{R}_+$; $n > m \in \mathbb{Z}_{++}$ implies $f(n) \geq f(m)$; and either $h(x) = x$ or $h(x) = g^{-1}(x)$. Each of these satisfies Aggregation.

## 4.4 Rank-dependent and critical-level generalized utilitarianism

Example 2 in the introduction showed that RDGU implies the unrestricted repugnant conclusion of Definition 2. Yet, neither maximin nor RDGU satisfies Aggregation. Similarly, critical-level generalized utilitarianism (Blackorby and Donaldson, 1984) satisfies Aggregation but does not satisfy either of the zero axioms unless it is "standardized" such that the critical level is zero[16]. All of these social orderings are addressed by Theorem 2, below. Here, we briefly note that the understanding in the literature that CLGU and RDGU

[15] Consider a totalist version: $W(\mathbf{u}) = \sum_{i=1}^{n(\mathbf{u})} \left(u_i\sqrt{n} - e^{-u_i} + 1\right)$. Then, $(-0.51, 0.1, 0.5) \prec 0\mathbf{1}_3$, but the order is reversed if both populations are replicated 100 times.

[16] Broome (2004) advances a case for CLGU which he "standardizes" by setting the critical level equal to zero, adjusting $g$ to match. 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 (n.d.) names "the weak repugnant conclusion," a further example of simultaneous debate in the prior literate about the *extent* and *acceptability* of the repugnant conclusion.

escape the repugnant conclusion depends on a *further* way in which the formalization of the repugnant conclusion varies in the literature.[17]

Each of Blackorby et al. (2005), Asheim and Zuber (2014), Arrhenius (n.d.) state a definition of repugnance in which $u^h > \varepsilon > 0$ and, implicitly, $n(\mathbf{v}) = 0$. However, Blackorby et al. (2005) and Asheim and Zuber (2014) further require that $n^\varepsilon > n^h$, but Arrhenius (n.d.) does not. In other words, they require that $n^\varepsilon$ be large. But it is even more repugnant to choose a lower quantity of lower-quality lives over a larger quantity of higher-quality lives. Additionally, Arrhenius (n.d.) interprets $\varepsilon$-lives qualitatively as "barely worth living," while Asheim and Zuber (2014) merely require that they be worse than $u^h$. For both RDGU (for any fixed $\beta$) and CLGU (for any fixed $c$) there exist cases where $x, y \in \mathbb{R}_{++}$, $x > y$, and $n(\mathbf{v}) = 0$ such that $y\mathbf{1}_m \succ x\mathbf{1}_n$ for some $m, n \in \mathbb{Z}_{++}$.[18] This, too, is a repugnant conclusion — and RDGU and CLGU imply it.

Moreover, although RDGU does not satisfy Aggregation, it implies the unrestricted very repugnant conclusion. The mechanism is the same as in Example 2. To see this, expand Definition 3 in a way that only intensifies the normative repugnance, by allowing the size of the high-utility population to be increased to any $\tilde{n}^h > n^h$. Then, choose a very-high-utility $\mathbf{v}$ and choose $\tilde{n}^h$ much larger than $n^\varepsilon$ and $n^\ell$.

Finally, CLGU implies its own very repugnant conclusion, which is proven analogously to the sadistic conclusion in the literature:

**Example 5.** *For any large $u^h \in \mathbb{R}_{++}$, any large $n^h \in \mathbb{Z}_{++}$, any very negative $u^\ell \in \mathbb{R}_{--}$, any large $n^\ell \in \mathbb{Z}_{++}$, and any small, positive $\varepsilon > 0$, there exists $\bar{\varepsilon} > \varepsilon$ and $n^{\bar{\varepsilon}}, n^\varepsilon \in \mathbb{Z}_{++}$ such that $u^\ell \mathbf{1}_{n^\ell} \cup \varepsilon \mathbf{1}_{n^\varepsilon} \succ u^h \mathbf{1}_{n^h} \cup \bar{\varepsilon} \mathbf{1}_{n^{\bar{\varepsilon}}}$.*

To see the proof, choose $\bar{\varepsilon} \in (\varepsilon, c)$ and make $n^{\bar{\varepsilon}}$ very large. Thus, CLGU chooses $u^\ell \mathbf{1}_{n^\ell} \cup \varepsilon \mathbf{1}_{n^\varepsilon}$ even though every person in it is worse off than every person in $u^h \mathbf{1}_{n^h} \cup \bar{\varepsilon} \mathbf{1}_{n^{\bar{\varepsilon}}}$; even though $u^h \mathbf{1}_{n^h} \cup \bar{\varepsilon} \mathbf{1}_{n^{\bar{\varepsilon}}}$ contains no negative-utility lives; and even though $u^\ell \mathbf{1}_{n^\ell} \cup \varepsilon \mathbf{1}_{n^\varepsilon}$ contains no excellent lives. This, too, is a very repugnant conclusion.

These examples offer further demonstrations 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 repugnant quantity-quality tradeoff. In some cases, indeed, the entailed repugnance is even more extreme — such as choosing a *smaller* worse-off population.

[17] To our knowledge, we are the first to note this discrepancy or its implications.

[18] 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 axiom 5).

# 5 Consequences of Non-Aggregation

Because same-number, risk-free[19] independence is an attractive axiom for evaluating economic policy, Theorem 1 applies to most social orderings that are used in welfare economics — even many that have commonly been interpreted to evade the repugnant conclusion. However, there are exceptions: Non-aggregative views such as maximin and other rank-dependent orderings. In this section, therefore, we broaden our scope to include orderings that take the opposite approach to social evaluation: those that satisfy Non-Aggregation. Non-Aggregation permits gains to enough of the worst-off members of the population to outweigh consequences for the better-off rest of the population.

**Axiom 10** (Non-Aggregation)**.** *For any* $\mathbf{u} \in \Omega$*, any* $\xi \in \mathbb{R}$*, and any* $\delta > 0$ *such that* $\xi + \delta < \min(\mathbf{u})$*, there exists* $n^* \in \mathbb{Z}_{++}$ *such that if* $n > n^*$*, then* $(\xi + \delta)\mathbf{1}_{n+n(\mathbf{u})} \succ \xi\mathbf{1}_n \cup \mathbf{u}$.

Non-Aggregation is satisfied by maximin, critical-level leximin (with positive or zero critical level, see Asheim and Zuber, 2014), and by RDGU (also with positive or zero critical level).

To extend to the broader set of social orderings that satisfy either Aggregation or Non-Aggregation, we must broaden the applied definition of the repugnant conclusion. 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** ($\varepsilon$-change)**.** Population $\mathbf{u}$ is separated only by an $\varepsilon$-change from $\mathbf{v}$ if either:

- $n(\mathbf{u}) = n(\mathbf{v}) + 1$ and $\mathbf{v} \cup \varepsilon\mathbf{1}_1 = \mathbf{u}$, 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 separated from another by two or more $\varepsilon$-changes, then any one person in the population may receive at most one $\varepsilon$-change.

[19] This paper does not put any restrictions on how $\succsim$ handles risky cases, because it only makes axioms about risk-free cases. Although a recent literature in welfare economics has debated the merits of separability in risky cases, there is wider agreement about risk-free cases.

With the definition of $\varepsilon$-change, we can define an extended very repugnant conclusion, of which the very repugnant conclusion is the strict subset of cases in which all $\varepsilon$-changes are additions of $\varepsilon\mathbf{1}_1$:

**Definition 6** (Extended very repugnant conclusion)**.** For any large $u^h \in \mathbb{R}_{++}$, any large $n^h \in \mathbb{Z}_{++}$, any very negative $u^\ell \in \mathbb{R}_{--}$, any large $n^\ell \in \mathbb{Z}_{++}$, and any small, positive $\varepsilon > 0$, there exists $n^\varepsilon \in \mathbb{Z}_{++}$, $m^\ell \geq n^\ell$, $m^h \geq n^h$, and $\mathbf{v}^\ell, \mathbf{v}^h \in \Omega$, $\mathbf{v} \in \Omega \cup \{\emptyset\}$ such that:

- $\mathbf{v}^\ell \succ \mathbf{v}^h$,
- $\mathbf{v}^h = u^h \mathbf{1}_{m^h} \cup \mathbf{v}$, and
- $\mathbf{v}^\ell$ is separated by $n^\varepsilon$ $\varepsilon$-changes from $u^\ell \mathbf{1}_{m^\ell} \cup \mathbf{v}$.

The extended very repugnant conclusion holds that many terrible lives full of suffering, which need never be lived, should be created, when many wonderful lives are available, merely so some other people receive tiny benefits. Every social ordering deemed plausible in the literature implies it:

**Theorem 2.** *If $\succsim$ satisfies the basic axioms (1-4), or otherwise has a $W(n, \Xi)$ reduced form representation with the properties in the Lemma, and satisfies either Aggregation (7) or Non-Aggregation (10), then $\succsim$ implies the extended very repugnant conclusion.*

*Proof.* See appendix section A.4. □

Theorem 2 implies that the extended very repugnant conclusion is implied by every social ordering that we are aware to be defended in the population economics literature.[20] It is implied by maximax. It is implied by Sider's (1991) Geometrism. It is implied by odd but imaginable examples such as:

- Rank populations by the sum of their two worst-off utilities, and rank populations of size 1 at 0.
- $V(\mathbf{u}) = \left(\frac{1}{n}\sum_{i=1}^{n(\mathbf{u})} g(u_i)\right) - \frac{\alpha}{n(\mathbf{u})}$, for $\alpha > 0$ (suggested for illustration by Partha Dasgupta).
- Negative utilitarianism: $V(\mathbf{u}) = \sum_{\{i:u_i<0\}} u_i$ (Smart, 1958).
- Any social ordering otherwise used in this paper, but lives of ranks that are not prime, divisible by three, or powers of ten are treated as though they do not exist.

[20] Theorem 2 below does not use these axioms of zero (5 and 6) because $\varepsilon$-changes need not involve lives near zero utility.

# 6 Conclusion

The contribution of this paper 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. 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.[21]

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 an instance of the repugnant conclusion. 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 equivalently repugnant cases.

The implication is that the repugnant conclusion cannot be escaped. Therefore, implying repugnance offers no 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." Although we do not argue that either aggregation or non-aggregation is tyrannical, we draw a similar conclusion for population ethics: axiomatic avoidance of a repugnant conclusion should be dropped as a methodological requirement for population economics.

Following this conclusion leaves open which family of social orderings to choose. We take no position on whether the social ordering should be averagist, totalist, prioritarian,

[21] Our axioms of zero 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.

utilitarian, or otherwise. 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 et al.'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 even deciding whether the critical level is zero or slightly positive. Or, we may simply rest assured that with $\beta$ close to 1, the policy recommendations of these two approaches will agree. 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 $\beta$ and $c$ and to alternative social orderings — each of which would imply repugnant conclusions in some imaginable case.

# A Proofs

## A.1 Proof of RDGU Examples 1 and 2

**Example 1.** The proof is by construction. If $g$ is not the identity function, use prioritarian-transformed utilities throughout the proof. Fix $\beta$ and any linear or concave continuous $g$ where $g(0) = 0$. $\gamma > 0$ is a very small utility increment which will be specified later. $\mu > r^*$ is a rank that will be set later. Then define utilities as follows:

$$\begin{array}{lll}
u_{[r]} = 1 - 3\gamma & v_{[r]} = 1 - 4\gamma & r = 1 \\
u_{[r]} = 1 - \gamma & v_{[r]} = 1 - 2\gamma & 1 < r < r^* \\
u_{[r]} = 1 & v_{[r]} = 1 + \varepsilon & r = r^* \\
u_{[r]} = 1 + \varepsilon + 2\gamma & v_{[r]} = 1 + \varepsilon + \gamma & r^* < r \leq \mu \\
u_{[r]} = 1 + \varepsilon + 3\gamma + \delta & v_{[r]} = 1 + \varepsilon + 3\gamma & \mu < r \leq \mu + m
\end{array}$$

By choice of a sufficiently large $\mu$, the large utility benefits of $\delta$ go to zero in RDGU's social value, because they become further and further in rank from $r^*$. By choice of a sufficiently small $\gamma$, the utility losses become small relative to $\varepsilon$.

For maximin and maximax, it is clear that the minimum and maximum are greater in $\mathbf{u}$. By choosing large enough $\mu$ and $\delta$ (presumably $\delta > \varepsilon$), the total utility in $\mathbf{u}$ can be made larger than the total utility in $\mathbf{v}$, including in $g$-transformation, which makes $\mathbf{u}$ preferred for any social ordering with generalized utilitarian same-number sub-principles (because $n(\mathbf{u}) = n(\mathbf{v})$), which includes total and average utilitarianism, prioritarian and egalitarian versions of these, critical-level generalized utilitarianism (Blackorby and Donaldson, 1984), and Ng's (1989) variable-value functional form.

**Example 2.** The essence of the proof is to make $\mathbf{v}$ very-high-utility and make $n^{\varepsilon} < n^{h}$. See Section 4.4 for a discussion of how alternative formalizations of the repugnant conclusion in the literature do or do not include cases where $n^{\varepsilon} < n^{h}$.

Let $n^{\varepsilon} = 1$. By hypothesis, $n^h$ is "large," which implies greater than one person. Choose $v \in \mathbb{R}_{++}$ such that $v > \frac{\sum_{i=1}^{n^h} u^h}{1-\beta^{n^h-1}} + 1$. Then let $\mathbf{v} = v\mathbf{1}_{10}$. The larger number of $u^h$ will cause RDGU to reduce the social value of the even-much-better lives in $\mathbf{v}$.

## A.2 Proof of RDGU Examples 3 and 4

The proof for Example 3 will construct a $\mathbf{v}$ with these properties. If $g$ is not the identity function, use transformed utilities. $\mathbf{v}$ has two levels of utility, $v^{\ell}$ and $v^h$, such that $0 < v^{\ell} < u^h < v^h$. Let there be 1 $v^h$ life, and choose a large number $n^{\ell}$ of $v^{\ell}$ lives such that the average utility of $\mathbf{v}$ is below $u^h$, so adding the $u^h$ lives increases the average (and total). Then, $\mathbf{v} \succ \mathbf{v} \cup u^h \mathbf{1}_{n^h}$ according to RDGU if

$$V(\mathbf{v}) = \beta\frac{1-\beta^{n^{\ell}}}{1-\beta}v^{\ell} + \beta^{n^{\ell}+1}v^h > \beta\frac{1-\beta^{n^{\ell}}}{1-\beta}v^{\ell} + \left[\text{value of } u^h \text{ lives}\right] + \beta^{n^h}\beta^{n^{\ell}+1}v^h = V\left(\mathbf{v} \cup u^h\mathbf{1}_{n^h}\right)$$

Which will be positive if

$$\left(1-\beta^{n^h}\right)\beta^{n^{\ell}+1}v^h > \left[\text{value of } u^h \text{ lives}\right]$$

Because the right hand side is finite, an arbitrarily large $v^h$ can be chosen to make the inequality true. To make the $u^h$ lives increase the average, it must be that $\frac{v^{\ell}n^{\ell}+v^h}{n^{\ell}+1} < u^h$. Let $v^{\ell}$ become close to zero from above. Then both inequalities can be met if $n^{\ell}$ is chosen such that $n^{\ell} + 1 > \frac{1}{1-\beta}$.

By continuity of RDGU, the proof can be extended immediately to Example 4 by choosing a sufficiently small $\varepsilon > 0$.

## A.3 Proof of Theorem 1

The basic axioms (1-4) are sufficient for an $n$-indexed family of $\Xi$ functions to exist, and for $W(n, \Xi)$ to be a reduced-form representation of $\succsim$ on $\Omega$.

The proof uses axiom 7 twice. Let $\mathbb{R} \ni \nu < \varepsilon$, and $\nu$ will be more specified later. The unaffected population will be constructed as $\mathbf{v} = \nu \mathbf{1}_m$ for some $m \geq 0$. By letting $m$ become large, Axiom 7 has the result that $\Xi\left(u^h \mathbf{1}_{n^h} \cup \nu \mathbf{1}_m\right)$ converges to $\nu$. Then, by letting $n^\varepsilon$ become large (much larger than $m$), $\Xi\left(\varepsilon \mathbf{1}_{n^\varepsilon} \cup u^\ell \mathbf{1}_{n^\ell} \cup \nu \mathbf{1}_m\right)$ converges to $\varepsilon$.

If $\succsim$ satisfies axiom 5, then choose $\nu \in (0, \varepsilon)$ and let $m$ become large and then the very repugnant conclusion is implied without the unaffected population having negative lives. If $\succsim$ satisfies axiom 6 but not axiom 5, choose $\nu < 0$.

## A.4 Proof of Theorem 2

Because the extended very repugnant conclusion implies the very repugnant conclusion, the proof of Theorem 1 applies for social orderings with a reduced form representation that satisfy Aggregation and a zero axiom.

For social orderings that satisfy Aggregation but not a zero axiom, or for social orderings that satisfy Non-Aggregation, the proof by construction uses the $u_j = v_j + \varepsilon$ horn of the definition of $\varepsilon$-change. For Aggregation axioms, simply include a very large base population, and in the combined population with bad lives, improve them all by $\varepsilon$.

To begin a construction for Non-Aggregation social orderings, choose any $\varepsilon$, $u^\ell$, $u^h$, $n^\ell$, and $n^h$ according to the EVRC. Next, set $m^h$ and $m^\ell$ in the EVRC to both be the maximum of $n^h + 1$ and $n^\ell + 1$. Then, let $\xi$ in the Axiom be $u^\ell - \varepsilon$ from the EVRC. Let $\delta$ in the axiom be $\varepsilon$ from the EVRC. Notice that $\xi + \delta$ in the Axiom now equals $u^\ell$ 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 Non-Aggregation axiom. We have now specified a $\mathbf{u}$, $\xi$, and $\delta$. So, there exists an $n^*$ such that if $n > n^*$ then $(\xi + \delta)\mathbf{1}_{n+n(\mathbf{u})} \succ \xi \mathbf{1}_n \cup \mathbf{u}$. 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 $\xi \mathbf{1}_{\tilde{n}} \cup \mathbf{u}$ from the Axiom is $\xi \mathbf{1}_{\tilde{n}} \cup u^h \mathbf{1}_{m^h}$, which is $\mathbf{v} \cup u^h \mathbf{1}_{m^h} = \mathbf{v}^h$ from the EVRC. Let $\mathbf{v}^\ell = (\xi + \delta)\mathbf{1}_{\tilde{n}+n(\mathbf{u})} = (\xi + \delta)\mathbf{1}_{\tilde{n}+m^\ell} = (u^\ell)\mathbf{1}_{\tilde{n}} \cup (u^\ell)\mathbf{1}_{m^\ell} = (\xi + \varepsilon)\mathbf{1}_{\tilde{n}} \cup (u^\ell)\mathbf{1}_{m^\ell}$. Set $n^\varepsilon$ equal to $\tilde{n}$. Finally, we can se that $(\xi + \varepsilon)\mathbf{1}_{\tilde{n}} \cup (u^\ell)\mathbf{1}_{m^\ell}$ is separated by $n^\varepsilon$ $\varepsilon$-changes from $\mathbf{v} \cup (u^\ell)\mathbf{1}_{m^\ell}$.

Maximin and maximax both have a reduced-form representation (simply as $\Xi$, with no sensitivity to $n$) and can be shown to imply the extended very repugnant conclusion by having **v** contain the least or greatest (respectively) utility level, and then increasing this with one $\varepsilon$-change.

# References


**Adler, Matthew D.** 2009. "Future Generations: A Prioritarian View." *George Washington Law Review*, 77(5/6): 1478.

**Anglin, Bill.** 1977. "The repugnant conclusion." *Canadian Journal of Philosophy*, 7(4): 745–754.

**Arrhenius, Gustaf.** 2000. "An impossibility theorem for welfarist axiologies." *Economics & Philosophy*, 16(2): 247–266.

——— 2003. "The very repugnant conclusion." *Uppsala Philosophical Studies*, 51.

——— n.d.. *Population Ethics: The Challege of Future Generations*: unpublished typescript (July 2017).

**Asheim, Geir B and Stéphane Zuber.** 2014. "Escaping the repugnant conclusion: Rank-discounted utilitarianism with variable population." *Theoretical Economics*, 9(3): 629–650.

——— 2018. "Rank-discounting as a resolution to a dilemma in population ethics." working paper, University of Oslo and PSE.

**Blackorby, Charles and David Donaldson.** 1984. "Social criteria for evaluating population change." *Journal of Public Economics*, 25(1-2): 13–33.

**Blackorby, Charles, Walter Bossert, and David Donaldson.** 1995. "Intertemporal population ethics: critical-level utilitarian principles." *Econometrica*: 1303–1320.

**Blackorby, Charles, Walter Bossert, and David J Donaldson.** 2005. *Population issues in social choice theory, welfare economics, and ethics*: Cambridge University Press.

**Blackorby, Charles, Daniel Primont, and R Robert Russell.** 1998. "Separability: A survey." *Handbook of utility theory*, 1: 49–92.

**Bossert, Walter.** 2017. "Anonymous welfarism, critical-level principles, and the repugnant and sadistic conclusions." working paper, University of Montréal.

**Boucekkine, Raouf and Giorgio Fabbri.** 2013. "Assessing Parfit's repugnant conclusion within a canonical endogenous growth set-up." *Journal of Population Economics*, 26(2): 751–767.

**Broome, John.** 2004. "Weighing lives." *Oxford*.

——— 2012. *Climate matters: Ethics in a warming world (Norton global ethics series)*: WW Norton & Company.

——— 2015. "Equality versus priority: a useful distinction." *Economics & Philosophy*, 31(2): 219–228.

**Budolfson, Mark and Dean Spears.** 2018. "Why the repugnant conclusion is inescapable." working paper.

**Cowen, Tyler.** 1996. "What do we Learn from the Repugnant Conclusion?" *Ethics*, 106(4): 754–775.

**Dasgupta, Partha.** 1995. "The population problem: theory and evidence." *Journal of economic literature*, 33(4): 1879–1902.

——— 2005. "Regarding Optimum Population." *Journal of Political Philosophy*, 13(4): 414–442.

**Fleurbaey, Marc.** 2010. "Assessing risky social situations." *Journal of Political Economy*, 118(4): 649–680.

**Fleurbaey, Marc and Bertil Tungodden.** 2010. "The tyranny of non-aggregation versus the tyranny of aggregation in social choices: a real dilemma." *Economic Theory*, 44(3): 399–414.

**Fleurbaey, Marc and Stéphane Zuber.** 2015. "Discounting, beyond utilitarianism." *Economics: The Open-Access, Open-Assessment E-Journal*, 9(2015-12): 1–52.

**Lawson, Nicholas and Dean Spears.** 2018. "Optimal population and exhaustible resource constraints." *Journal of Population Economics*, 31(1): 295–335.

**Nerlove, Marc, Assaf Razin, and Efraim Sadka.** 1982. "Population size and the social welfare functions of Bentham and Mill." *Economics Letters*, 10(1-2): 61–64.

**Ng, Yew-Kwang.** 1989. "What Should We Do About Future Generations?: Impossibility of Parfit's Theory X." *Economics & Philosophy*, 5(2): 235–253.

**Palivos, Theodore and Chong K Yip.** 1993. "Optimal population size and endogenous growth." *Economics letters*, 41(1): 107–110.

**Parfit, Derek.** 1984. *Reasons and Persons*: Oxford.

——— 2016. "Can we avoid the repugnant conclusion?" *Theoria*, 82(2): 110–127.

**Pivato, Marcus.** 2018. "Rank-additive population axiology." conference paper.

**Roberts, Melinda A.** 2011. "The Asymmetry: A Solution." *Theoria*, 77: 333–367.

**Scovronick, Noah, Mark B Budolfson, Francis Dennig, Marc Fleurbaey, Asher Siebert, Robert H Socolow, Dean Spears, and Fabian Wagner.** 2017. "Impact of population growth and population ethics on climate change mitigation policy." *Proceedings of the National Academy of Sciences*, 114(46): 12338–12343.

**Sider, Theodore R.** 1991. "Might theory X be a theory of diminishing marginal value?" *Analysis*, 51(4): 265–271.

**Smart, Roderick Ninian.** 1958. "Negative utilitarianism." *Mind*, 67(268): 542–543.

**Spears, Dean.** 2017. "Making people happy or making happy people? Questionnaire-experimental studies of population ethics and policy." *Social Choice and Welfare*: 1–25.

**Temkin, Larry S.** 2014. *Rethinking the good: Moral ideals and the nature of practical reasoning*: Oxford University Press.

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