Does the Repugnant Conclusion have important implications for axiology or for public policy?

For Campbell, Bykvist, and Arrhenius eds. *Oxford Handbook of Population Ethics*

August 2019

8,997 words

Mark Budolfson

Harvard, ANU, UVM

Dean Spears

UT-Austin, ISI-Delhi, IZA, IFFS

Abstract

Formal arguments have proven that avoiding the Repugnant Conclusion is impossible without rejecting one or more highly plausible population principles. To many, such proofs establish not only a deep challenge for axiology, but also pose an important practical problem of how policymaking can confidently proceed without resolving any of the central questions of population ethics. Here we offer deflationary responses: first to the practical challenge, and then to the more fundamental challenge for axiology. Regarding the practical challenge, we provide an overview of recent literature that explores the implications for public policy of the Repugnant Conclusion and related puzzles within population ethics, and shows that there is more commonality in the implications of different population axiologies than is generally recognized. The upshot is that uncertainty about population ethics presents no important problem for policymaking, and the importance of population ethics to policymaking has been overstated. We then turn to more fundamental issues about axiology, and describe a new series of formal proofs that undermine the idea that *any* plausible axiology could avoid the Repugnant Conclusion, including those that are assumed to avoid it in the literature. The philosophical upshot is that the Repugnant Conclusion cannot plausibly be avoided, and so it is a mistake to assume as a constraint on a plausible axiology that it must be avoided, as is assumed by most of the literature in population ethics, including by all of the fundamental impossibility theorems.

Keywords

population ethics, population axiology, repugnant conclusion, value theory, demography, climate change, consequentialism, utilitarianism, policy, feasibility

1. Introduction

Population ethics is a field at the intersection of economics and ethics. A central question within population ethics is what is the correct *population axiology*, which ranks outcomes composed of different numbers of individuals with different levels of well-being. A key challenge within population axiology is how to rank outcomes in which populations differ in tradeoffs between quantity and quality – i.e., how should outcomes that create more lives be weighed against outcomes that create fewer but better lives? For example, consider two possible outcomes: Population A, a population of ten billion people each of whom has the same very high quality of life, versus Population Z, a far larger population composed entirely of people whose lives are all barely worth living; assume also that other things are equal between A and Z, and thus that there are no other normatively relevant differences between A and Z.[1] If a population axiology implies that for a sufficiently large population in Z, that Z is better than A, then that axiology implies the **Repugnant Conclusion**.

For several decades, the population ethics literature has focused on the Repugnant Conclusion and the effort to devise a suitable axiology that escapes it (Parfit, 1984).[2] Figure 1 illustrates how an axiology might imply an instance of the Repugnant Conclusion, and how an axiology might avoid implying it:

[1] In what follows, we will always implicitly understand this *other things are equal* condition as understood in our discussion of specific cases.
[2] In contrast, a few philosophers have argued that we should endorse the Repugnant Conclusion – we outline their arguments in Section 5 below.

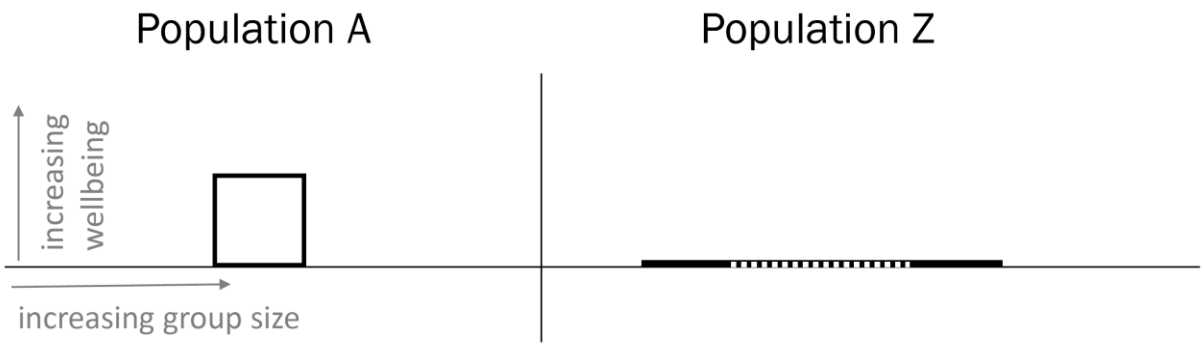


Figure 1: An Instance of the Repugnant Conclusion for Total Utilitarianism: Population A consists of 10 billion equally high quality lives, Population Z consists of a sufficiently large population of equally barely worth living lives so that the sum total of wellbeing in Z is larger than in A[3]

Total Utilitarianism, which ranks populations according to their total sum of wellbeing, implies the Repugnant Conclusion by preferring Z to A in Figure 1; Average Utilitarianism, which ranks populations according to average wellbeing, avoids this specific instance by preferring A to Z, but has other undesirable implications.

Formal arguments by Ng (1989), Arrhenius (n.d.), and others have proven that avoiding the Repugnant Conclusion is impossible without rejecting one or more other highly plausible population principles. To many, such proofs establish not only a deep challenge for axiology, but also pose an important practical problem: how can climate policy, population policy, or human development policy confidently proceed without resolving any of the central questions of population ethics, such as whether and how to avoid the Repugnant Conclusion?

Here we offer deflationary responses: first to the practical challenge, and then to the more fundamental challenge for axiology. Regarding the practical challenge, we argue that the importance of population ethics to realistic decision-making has been overstated. We provide an overview of recent literature that explores the implications for public policy of the Repugnant Conclusion and related puzzles within population ethics, and shows that there is more agreement in the implications of different

[3] In figures of this type, used throughout this chapter, each box represents a group of people. The height of the box is proportional to equal per-person well-being of the group and the width is proportional to number of people; boxes are to scale unless noted otherwise.

population axiologies than has previously been recognized. The upshot is that uncertainty about population ethics presents no important problem for policymaking.

We then turn to more fundamental issues about axiology, and describe a new series of formal proofs that undermine the idea that *any* plausible axiology could avoid the Repugnant Conclusion, including those that are assumed to avoid it in the literature. The philosophical upshot is that the Repugnant Conclusion *cannot* plausibly be avoided, and so it is a mistake to assume as a constraint on a plausible axiology that it *must* be avoided, as is assumed by most of the literature in population ethics, including by all of the fundamental impossibility theorems.

A unifying theme of both the practical and philosophical discussions is that there is more commonality in the implications of different population axiologies than has previously been recognized: in part this is because of empirical facts such as the improving trajectory of average wellbeing, in part this is because policy choices are coarse and constrained by feasibility, in part this is because just as policies are already evaluated for their suitability under alternative values of a statistical life or social discount rate, so too policy can be evaluated for acceptability under a range of population axiologies, and in part this is because the philosophical literature has made an important mistake by focusing only on a proper subset of instances of the Repugnant Conclusion, which has supported the mistaken idea that avoiding the Repugnant Conclusion should be a constraint on a plausible axiology.

2. Empirical demography and feasibility constraints on timely, non-coercive population reduction

The 2014 International Panel on Climate Change report expressed the worry that climate policy might depend importantly on population ethics, and might therefore inherit the uncertainty of population axiology. The connections between population and climate take several forms. On the one hand, increases in the size of the population will lead to increases in carbon emissions; so, policy to reduce population growth might be a useful form of climate mitigation policy. On the other hand, climate change may change the size of the future population (such as by increasing early-life mortality; Geruso and Spears, 2018), so population implications may change the social desirability of emissions-reduction policy. Either way, knowing what climate

policy to pursue may seem to require knowing how to value increases or decreases in the size of the population. So, it would appear that climate change provides a clear example of how making progress on the central disagreements in population ethics is a prerequisite for confident policymaking, given the radical disagreements that are possible between axiologies like averagism and totalism when it comes to valuing changes in the size of the population (Broome, 2012).[4]

In contrast, we argue that population ethics is unlikely to have much significance for policy, relying on the example of climate policy as a useful illustration. First, in this section, we note that even very ambitious population policy would be unlikely to substantially change the size of the population in the decades-long short-term that is relevant for climate mitigation policy. Climate mitigation is urgent, and in the meanwhile technological progress may rapidly change the environmental costs of increased population size. In the next section, we show that even if climate change will change the size of the future population, all plausible axiologies are likely to consider decarbonization to be a policy goal as quickly as is likely to be politically feasible. These examples illustrate a more general point that uncertainty over axiology need not translate into uncertainty about policymaking.

Escaping the worst of climate change will require rapid reductions in carbon emissions over the coming decades in the first half of the 21st century. Here, *timing* is important: achieving the goals of the Paris Agreement, for example, would require carbon emissions to be essentially eliminated worldwide within the length of an average human lifespan, and to be substantially reduced even more quickly. Some people have proposed that reducing the growth of the human population would be a useful policy strategy to achieve reduction in carbon emissions. The argument is that it is an empirical fact that larger populations emit more polluting carbon all else equal (O'Neill, *et al.*, 2012), so reducing population growth would appear to offer an effective strategy to reduce population emissions. However, populations change slowly over this

[4] "We do not know what value to set on changes in the world's population. If the population shrinks as a result of climate change, we do not know how to evaluate that change. Yet we have reason to think that changes in population may be one of the most morally significant effects of climate change. The small chance of catastrophe may be a major component in the expected value of harm caused by climate change, and the loss of population may be a major component of the badness of catastrophe. … So we face a particularly intractable problem of uncertainty, which prevents us from working out what we should do. Yet we have to act; climate change will not wait while we sort ourselves out" (Broome (2012), pg. 183-185).

timescale, and in light of this Budolfson and Spears 2019a observe that this argument overlooks two important facts from empirical demography. The first is that population growth is already slowing. The size of the population is projected to peak around 2100 (Gerland, *et al.* 2014). This is because the world is progressing through the demographic transition, characterized by paradigm-shifting improvements in the developing world in health and human development, such as sanitation reduced mortality rates (Deaton 2013; Coffey and Spears, 2017), leading to a subsequent fall in fertility rates. Because mortality rates have a lower bound near zero and fertility rates can also only fall so low, the uniquely rapid population growth of the mid-20th century — when mortality rates had fallen but fertility rates had not yet followed down — could not ever be projected to be repeated (Lam, 2011). So, any fertility policy to reduce the size of the population would have the task of *accelerating* a decline in growth rates that is already underway.

The second fact is that even much-larger-than-feasible changes in fertility and mortality rates would not substantially change the size of the population over the decades-long medium-term that is relevant for climate mitigation. This is because of *population momentum*: demographers' term for changes in the size of the population that would continue to occur due to the age distribution of the population, even if fertility and mortality rates hypothetically changed to replacement levels. In short, in high-fertility countries, there are currently more young girls at each age than adult women, so their inevitable aging guarantees continued increase in the size of birth cohorts over the coming decades.

Figure 2: Two population projections are similar for coming decades

Source: UN Population Division projections

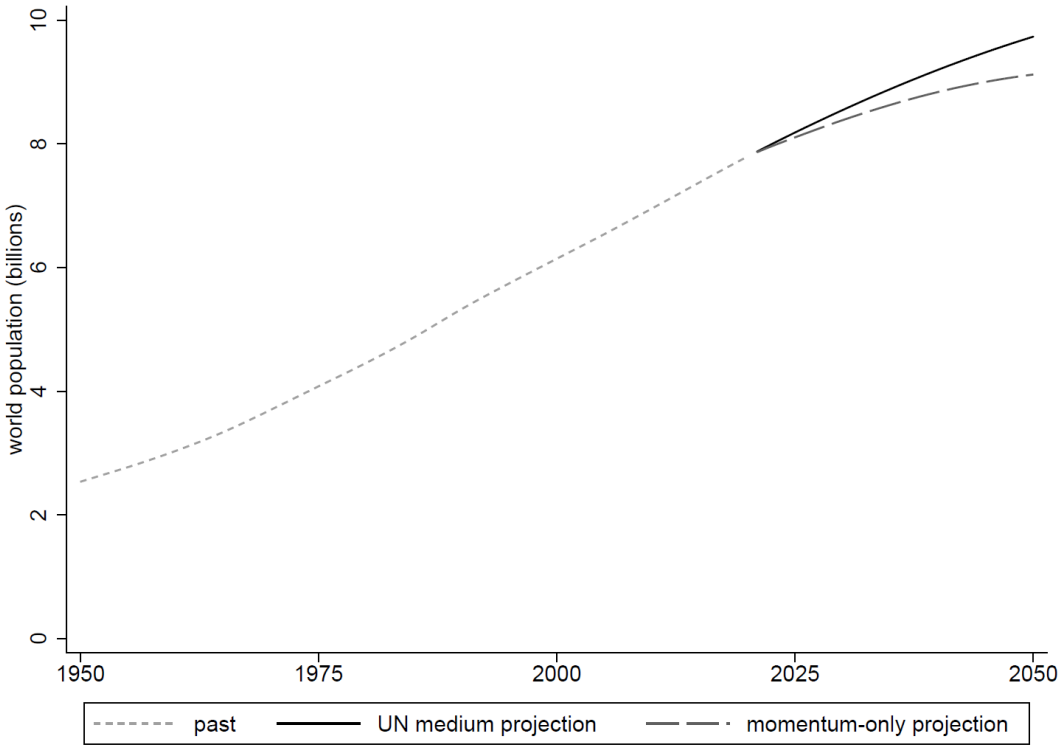


Figure 2 plots the UN's projection of future population size under two scenarios: the most likely path and a population momentum path. The population momentum path projects what would happen if mortality and fertility rates instantly went to replacement levels – i.e. went to replacement levels starting *today*. We emphasize that the momentum path is not a policy option; it is a highly counterfactual illustration of population growth falling much faster than is actually feasible: no actual fertility policy implementable by states could plausibly cause such a large, rapid change in fertility rates. Our purpose in showing Figure 2 is to demonstrate that, over the coming decades relevant for climate mitigation policy, the size of the population is bounded for policymaking purposes, as even under the much-greater-reductions-than-are-feasible population momentum path there is a surprisingly small reduction in population over the coming decades, which would translate into only a small reduction in greenhouse gas emissions relative to the reductions that are needed. To see this, note that by 2050, the population would be about 9 billion people under the momentum-only path and 10 billion under the medium variant; under the O'Neill, *et*

al. (2012) elasticity estimate, this unfeasibly large population change would imply only a 10% reduction in emissions from where they would otherwise be, at a time when many climate scientists argue emissions should be approaching zero. As time continues, the difference between the two projections becomes larger, but 2100 is already late in the game relative to both the policy horizon for climate mitigation and the plausible emergence of "backstop" low-emissions technology, given that the key to any decent solution to climate change is large-scale decarbonization of the economy over the next roughly fifty years.

One upshot is that it is not urgent to normatively evaluate the goodness or badness of differences in population size for climate policymaking, because if the size of the future population cannot be reduced over a timescale that matters to climate mitigation policy, then this eliminates one reason for thinking that the puzzles of population ethics are practically urgent.

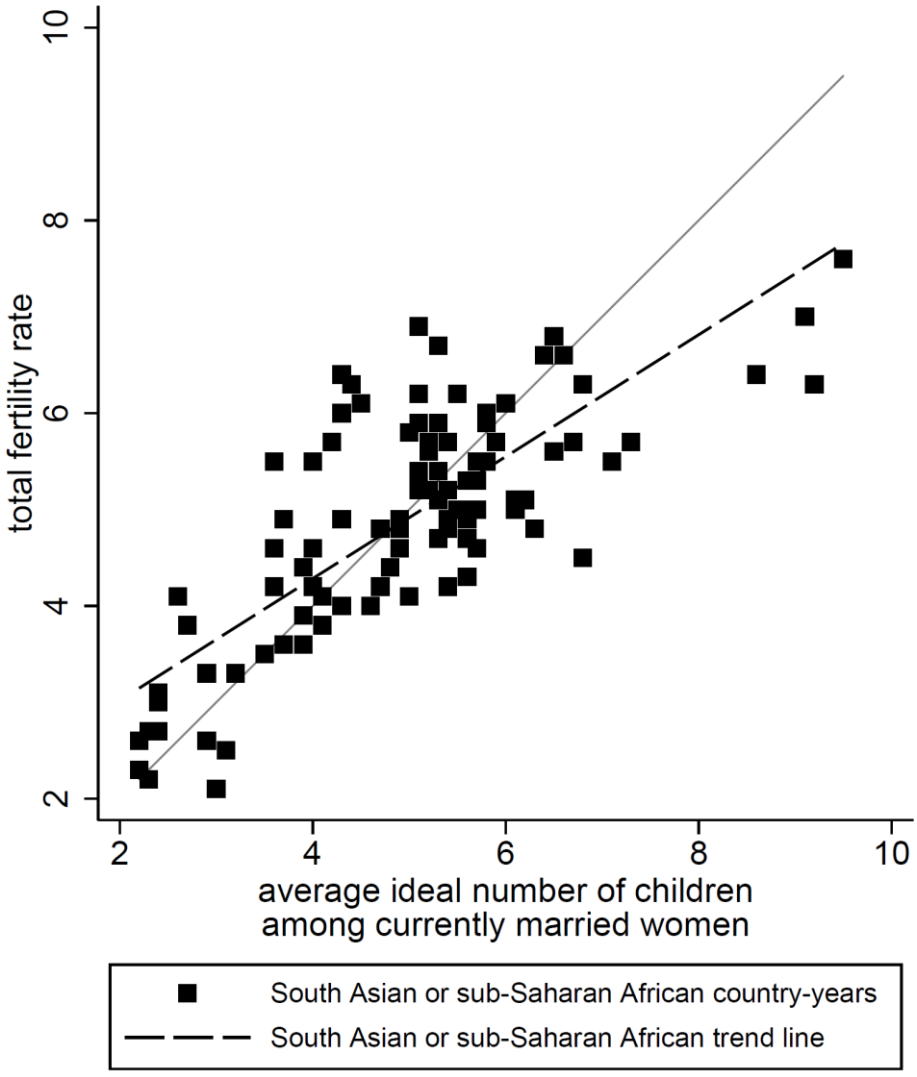
A further lesson from history and demography is that policy-making to change fertility rates is neither easy nor costless in practice. Indeed, the actual history of population policies implemented by states, especially in developing countries, has involved considerable coercion and harm is some cases, as well as benefits and freedom in other contexts (Connelly, 2009). With this in mind, it is important to consider the argument that population reduction is generally an undesirable policy lever – especially in the context of global problems such as climate change – because insofar as developed nations try to incentivize *non*-coercive population reduction in developing nations, the expected effect would be that (as we have observed historically) developing nations instead adopt *coercive* measures to achieve the relevant targets, and the result is that the costs of the policy greatly exceed the benefits, including in terms of human rights violations.

Another reason that it can be challenging for policy to accelerate the ongoing decline in fertility rates is that *achieved* total fertility rates often reflect *desired* total fertility rates. In other words, on average, population-level total fertility rates substantially reflect preferences. Some pregnancies are unintended, but over a childbearing career a woman who experiences an unintended pregnancy can nevertheless achieve her desired total number of children (a quantity known to population science as "ideal fertility") by reducing her subsequent fertility (such as by

increasing subsequent breastfeeding or other contraceptive effort to space or avoid later pregnancies). If so, unexpected pregnancy would influence the timing and spacing of fertility, but would only have a small effect on population-level average total fertility.

Figure 3: The correlation across populations between ideal and achieved fertility

Source: Authors' use of Demographic and Health Survey data to update Pritchett (1994)



The empirical demography literature has long studied the important relationships between fertility intentions and fertility behavior. A classic analysis is Pritchett's

(1994) examination of the results of fertility surveys, which we have updated with more recent data in Figure 3. Comparing across developing countries (including populations in places and times where access to modern contraceptive technology was much more limited than today), Figure 3 shows that ideal total fertility rates are highly correlated with achieved total fertility rates. In other words, high fertility rates often happen because they are wanted.

An useful example for our purposes is the large population science literature on the extreme case of the one-child policy in China: how much of China's late-20th-century fertility decline was due to this policy and its implementation, rather than to economic and social changes that were also happening, at different paces, around the world? (Greenhalgh 2008). For example, a 2018 issue of the journal *Demography* contains a heated debate over this issue, with three detailed, published responses criticizing an earlier article that claimed that the one-child policy caused total fertility rates to be half-a-birth-per-woman lower than they otherwise would have been.[5] In short, many demographers who are critical of the alleged influence of the one-child policy argue that social and economic changes, qualitatively similar to trends in other developing countries, were principally responsible for China's fertility decline, operating importantly through changes in ideal fertility (Cai 2010). Babiarz, *et al.* (2019) have noted, for example, that China's fertility decline substantially *precedes* the 1980 introduction of the one-child policy: the total fertility rate fell from around 6 in 1970 to below 4 in 1980, which as Figure 3 shows was not aberrant in international comparison. The upshot for our argument is that the population science literature is far from agreed that the one-child policy had a large effect on fertility rates. In addition, no leading figure in the climate policy debate would seriously advocate a policy so coercive as a form of climate mitigation.

Around the world, fertility rates today are lower than Pritchett documented them to be in the 1990s and 1980s. What today are the highest total fertility rates are around 5 children per woman's childbearing career, in some sub-Saharan African countries. But total fertility in India, for example, is now about 2.3, and in many Latin American countries the total fertility rate has fallen below replacement. Of course, it is

[5] Goodkind 2017, the paper at the center of the recent debate, drew its conclusions by comparing China to Vietnam as a counterfactual. For a response based on the details and timing of China's fertility patterns, see Zhao and Zhang 2018.

possible that these rates could fall even further. The important point in this context is that achieving much lower population-average total fertility rates is unlikely to be as simple as further increasing access to modern contraception, in part because investments against the HIV/AIDS epidemic have already done much to increase access to reproductive health care. Instead, if high fertility rates tend to reflect, at least in part, high fertility preferences, what policy can achieve without coercion may be limited to long-term changes in fertility preferences through education and human development. Such policy may be an important component of humanity's long-term environmental strategy, but is unlikely to make a large difference to carbon emissions in the near-term decades relevant to climate mitigation, given that the key to any decent solution to climate change is large-scale decarbonization of the economy over the next roughly fifty years.

The upshot is that population momentum and other considerations create stringent feasibility constraints on the magnitude of possible non-coercive population reduction in the near term. For policy challenges like climate change that require large changes within a few decades, population reduction is therefore not well suited to serve as an important policy lever, and so uncertainty about population ethics presents no important challenge for policymaking.

3. Agreement across axiologies: convergence on a corner solution

Section 2 considered the possibility that fertility-reduction policy could be used as a form of climate mitigation policy. This section considers another possible interaction between population and climate policy. One of the possible consequences of climate change is changing the size of the future population, for example because of increasing mortality rates. If so, it may be that choosing the correct carbon mitigation policy (such as choosing an optimal tax on carbon emissions) depends on the population axiology, because the social evaluation of climate damages depends on the social evaluation of the change in the size of the population (Broome 2012).

Arrhenius *et al.* 2019 argue that this worry is likely to turn out not to be important in practice. The reason is that the policy-relevant set of possible climate mitigation policies (such as e.g. the optimal 2150 decarbonization rate) is both

bounded and coarse. Being *bounded* means that decarbonization cannot exceed some increasing amount each year due to political and other feasibility constraints. For example, it is likely that it is not politically feasible for the world to do anywhere near 100% of the mitigation that optimal policy models say we should be doing in the short and medium term. Being *coarse* means that feasibility constraints may also limit the fineness of the plausible distinctions between climate policies. For example, it may be that many political actors can only choose a broad policy such as "vigorously pursue multisectoral, rapid decarbonization as an urgent policy priority." Given these constraints, it is likely that all axiologies agree on the same climate policies into the future within this feasible set: roughly, to decarbonize as quickly as is politically feasible. In the terminology of optimization in calculus, this is called a *corner solution*: the optimum is a unique point at the boundary of the feasible set. Whether or not these features in fact describe the climate policy problem is an empirical question.

If the optimal climate mitigation policy is a corner solution, then it might not matter to policymaking what the correct population axiology is. Consider the illustrative case of average and total utilitarianism, commonly taken to be theoretical opposites in the population ethics literature. Average utilitarianism seeks to maximize average wellbeing; total utilitarianism seeks to maximize the total sum of wellbeing, the product of average wellbeing and the population size. It is plausible that climate damages will both decrease average wellbeing and will decrease the future population size, relative to a future in which there were no climate damages. Even ignoring the effect on population size, the effect of climate change on average wellbeing is sufficient to justify a corner solution in mitigation policy: "pursue full decarbonization in the next 50 years." Now imagine that a policy-maker realizes that instead of maximizing average wellbeing, she should have been maximizing total wellbeing. She was already pursuing the most aggressive available mitigation policy. So, the additional consideration of population size does not change the optimal action available in her choice set.

Of course, it is possible that mitigation policy-making is not so coarse. If so, different axiologies might disagree on details, such as exactly how fast to pursue decarbonizaiton, or exactly what industries to prioritize, even if they agree on the broad outline of climate policy. It is also the case that different axiologies could agree

on the optimal action while disagreeing about the comparative goodness and badness of options: for example, totalism could think that 90% decarbonization is much worse than 100%, while averaging thinks 90% decarbonization is only a little worse than 100%. Figure 4 illustrates this possibility with a hypothetical example. Our point is merely that if optimal feasible policy turns out to be a corner solution in a coarse set of politically feasible options, then a wide range of divergent axiologies may turn out to agree on what policymakers should do.

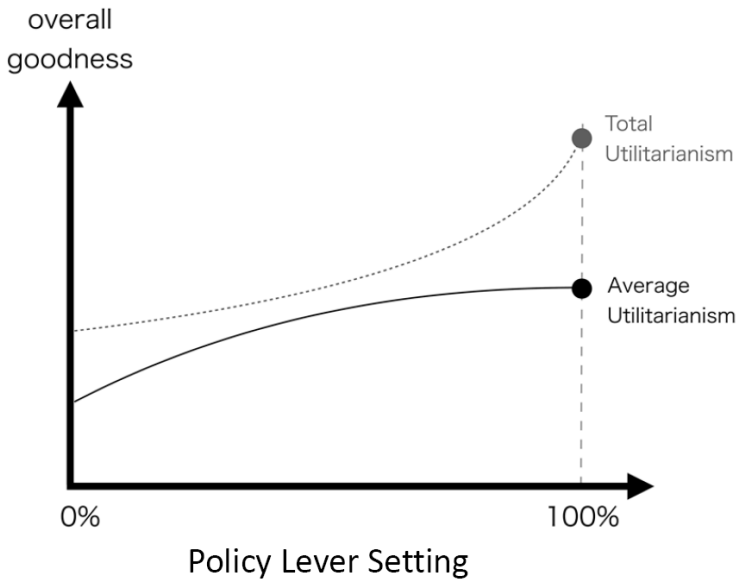


Figure 4: Total and Average Utilitarianism agree on what choice is best, while disagreeing on the comparative value of outcomes

Although Figure 4 is hypothetical and stylized, Scovronick et al. 2017 provide detailed quantitative computations that suggest that the actual case of setting optimal climate policy may have this corner solution property. They compute optimal carbon taxes for intertemporal versions of Total and Average Utilitarianism. In both cases, the recommended mitigation policy is more ambitious that is currently being pursued, and very likely more than is politically feasible.

4. Agreement across axiologies: convergence given increasing wellbeing

So far, this chapter has illustrated the implications of population axiologies for policy via the example of climate policy. This section turns to policymaking outside of climate change — for example, policy intended to decrease or increase fertility rates. In short, should policy-makers seek to increase or decrease the size of the future population, or is the size of the future population irrelevant for overall social goodness?

This might initially appear to be a question on which population axiologies such as average and total utilitarianism disagree. Total utilitarianism, the conventional wisdom holds, would favor increasing the number of future well-off lives, while average utilitarianism does not seek an increase in the size of the population, separate and apart from any second-order effects on average wellbeing. Spears and Budolfson 2019a observe that this conventional view is false insofar as it overlooks the fact that these population axiologies rank the goodness of the full intertemporal population: past, present, and future.

In particular, Spears and Budolfson build upon an empirical premise: average well-being will be greater among people in the future than among people in the past, and should be expected to continue increasing over time. They summarize the demographic and statistical evidence for this empirical premise, rooted in the important facts of the epidemiological transition from a past world of high mortality rates to a contemporary world of low mortality, low poverty, and high literacy (Rosling, 2018; Deaton, 2013; Coffey and Spears, 2017). More formally, this empirical premise can also be generalized and stated in axiologically-neutral terms of the "equally-distributed equivalent" level of well-being of a population, which allows for an axiologically neutral summary of well-being (Fleurbaey, 2010).

> **Equally-distributed equivalent.** For any population A, the *equally-distributed equivalent* of A, according to a specific axiology, is the level of well-being u such that a population the same size as A in which everyone was at u would be ranked exactly as good as A by that axiology.

The concept of equally-distributed equivalent generalizes the role of average well-being in a utilitarian axiology: it captures how well-off a population is, for a population of its

size, according to a specific axiology. In this way, it is like the body mass index, which allows weight to be compared across people of different height. For any version of utilitarianism, the equally-distributed equivalent is simply average wellbeing, but for other axiologies (such as versions of prioritarianism, maximin, or rank-dependent views) the equally-distributed equivalent can differ from the simple average, in a way that derives from the different properties of the axiology. With this definition, we can state the Empirical Premise:

> **Empirical Premise.** The future will be better than the past: more precisely, the equally-distributed equivalent well-being of the future will be greater than the equally-distributed equivalent of the past (and will be greater than the critical level, if an axiology has one).

The next step is to combine the Empirical Premise with a broadly attractive axiom for population ethics, Convergence in Equivalence, which for generality and neutrality between competing axiologies uses the concept of the equally-distributed equivalent. Convergence in Equivalence is accepted by a wide range of population axiologies, including total, average, variable-value, and critical-level versions of utilitarianism, prioritarianism, and egalitarianism, among many other axiologies:

> **Convergence in Equivalence.** Starting from any population A, adding a very large number of lives at an wellbeing level u, as the number of added lives become large, the equally-distributed equivalent converges to u.

To reject Convergence in Equivalence is to hold that an arbitrarily tiny fraction of the population can be more important for social evaluation than the vast majority of the population. Convergence in Equivalence would not be accepted by some rank-dependent views, nor by Non-Archimedean axiologies or those that endorse Weak Value Superiority (Parfit, 2016), such as presented by Erik Carlson in this volume; we discuss these views in the next section, and note that they have implications at least as counterintuitive as the repugnant conclusion.

Convergence in Equivalence, combined with the Empirical Premise, in the context of some formal properties[6] that are generally uncontroversial even among axiologists who otherwise disagree, implies that it could be an improvement to increase the size of the future population, even in a way that makes all future people worse off than they otherwise would be. More formally, relative to the business-as-usual size n and equally-distributed-equivalent wellbeing u of the future population, there exists $m > n$ and $v < u$ such that a population with (m, v) would be strictly better than a population with (n, u).

The intuition behind this result can be illustrated using the case of average utilitarianism, merely because average utilitarianism provides a conceptually simple illustration. Average utilitarianism wishes to maximize the average well-being among all past, present, and future people who will ever have lived. If future people will be better-off than past people, then a change that vastly increases the number of future people while slightly decreasing their well-being would increase the overall, intertemporal average.[7] Of course, it may be that another possible future would be even better: this is merely a claim relative to the business as usual. Different axiologies have different degrees of overlap in the future paths that they prefer to business as usual.

Many past and present states articulate or implement *practical population policy*, by which we mean programs or efforts to influence population-level fertility rates, such as contraception or sterilization programs, in the anti-natalist direction, or child care subsidies and parental leave policies, in the pro-natalist direction (Connelly, 2009). Given the result just noted, it is unwarranted to assume that different axiologies such as average and total utilitarianism would have radically different implications for practical population policy. We speculate that this unwarranted assumption is nevertheless widespread partly because textbook examples of population ethics often ignore the fact that practical population policy concerns future

[6] The formal properties are "Egalitarian replication is not worse at high enough wellbeing," "Same-number egalitarian Pareto," and "Same-number egalitarian continuity." These are formally defined in Spears and Budolfson 2019a.

[7] Although the population axiology literature accepts that population axiologies rank the full set of all lives (past, present, and future), some readers may be surprised that the past matters. Spears and Budolfson 2019a detail preference-reversal reasons that may account for why the population ethics literature has taking this as part of the definition of a population axiology (Arrhenius, n.d.).

additions to the set of past and present lives which have already come to exist. In other words, the practical implications of population ethics are not about comparisons of non-overlapping populations that are merely metaphysically possible and entirely distinct, but rather are about comparisons of overlapping populations that include an unaffected base population (namely, past people and unaffected current people, who are worse off on average than future people will be). As the next section will explore, the possibility of an unaffected base population is not merely practically important, but is also theoretically important for population ethics.

5. Why the Repugnant Conclusion should not constrain the search for the correct population axiology

This section considers a final reason why some puzzles of population ethics might not pose a problem: namely, that there may be compelling arguments against a methodological requirement to avoid the Repugnant Conclusion, which would then disarm the main challenge posed by the impossibility results within population axiology, since the impossibility results all assume that the Repugnant Conclusion must be avoided.

One familiar argument for a deflationary response insists we should *endorse* the Repugnant Conclusion *even if it can be avoided by axiologies that are otherwise attractive*. The most straightforward arguments of this type are offered by Tännsjö 2002, Huemer 2008, and others: namely, that that low-positive-value lives should be recognized as socially valuable, and therefore welfarist axiology that does not do so should be off the table. This kind of argument is sometimes supplemented by the additional claim that we are misled by our intuitions about the Repugnant Conclusion because we are unable to fully appreciate how much value there is in a population with many billions of lives that are barely worth living, because we are in general unable to have reliable intuitions about cases involving very large numbers (Tännsjö 2002, Broome 2004, see also Pummer 2013 for discussion). A related argument sometimes offered for endorsing the Repugnant Conclusion is that the conjunction of the principles that imply the Repugnant Conclusion is more compelling than is the intuition that the Repugnant Conclusion should be avoided. For example, several recent papers have shown that incorporation of probability and risk into population

ethics point towards an additive or generalized totalist framework, with Repugnant-Conclusion-type implications (Arrhenius and Stefánsson 2018, Nebel 2019, Spears 2018; see also Broome 2004). So, any repugnance of the conclusion would be set against the attractiveness of rational decision-making under risk.

In what follows we highlight a different kind of argument against the idea that the Repugnant Conclusion must be avoided: namely, that the Repugnant Conclusion should not be avoided *because it cannot be avoided by any generalized utilitarian axiology*, or for that matter by any other axiology that we should find plausible. Here and in what follows, "generalized utilitarian axiology" is used as formally defined by Blackorby, et al (2005) and includes all total, average, critical-level, and variable-value versions of utilitarianism, prioritarianism, and egalitarianism – not merely utilitarianism.[8] The moving parts of this argument involve, first, the observation that there are instances of the Repugnant Conclusion that cannot be avoided by average utilitarianism, and that cannot be avoided by many other axiologies that are commonly assumed to avoid the Repugnant Conclusion – so, the Repugnant Conclusion is an implication of more axiologies than is commonly recognized; second, the articulation of a sufficient condition for being an instance of the Repugnant Conclusion that includes these instances, which are generally ignored in more simplistic formal definitions of the Repugnant Conclusion in the literature; and finally, a formal proof that exploits the improved characterization of the Repugnant Conclusion to prove that in fact *all* generalized utilitarian axiologies (and others) imply the Repugnant Conclusion when it is correctly characterized. This argument can be supplemented with an additional analysis of what the Repugnant Conclusion is most fundamentally, which then allows for more precise arguments that no plausible axiology could hope to avoid it, up to and including additional formal proofs. In the rest of this section, we explain these arguments in more detail, following their initial presentation in the philosophy literature by Budolfson and Spears 2018 and in the economics literature by Spears and Budolfson 2019b.

We begin with the observation that some instances of the Repugnant Conclusion cannot be avoided by average utilitarianism and other axiologies that are commonly assumed to avoid the Repugnant Conclusion. For example, consider the

[8] This set of axiologies is known as "generalized utilitarian" in the economics literature because they share a key property of utilitarianism: same-number anonymous additive separability.

three different choices in Figure 5, which illustrate three different kinds of choices where instances of the Repugnant Conclusion can arise:

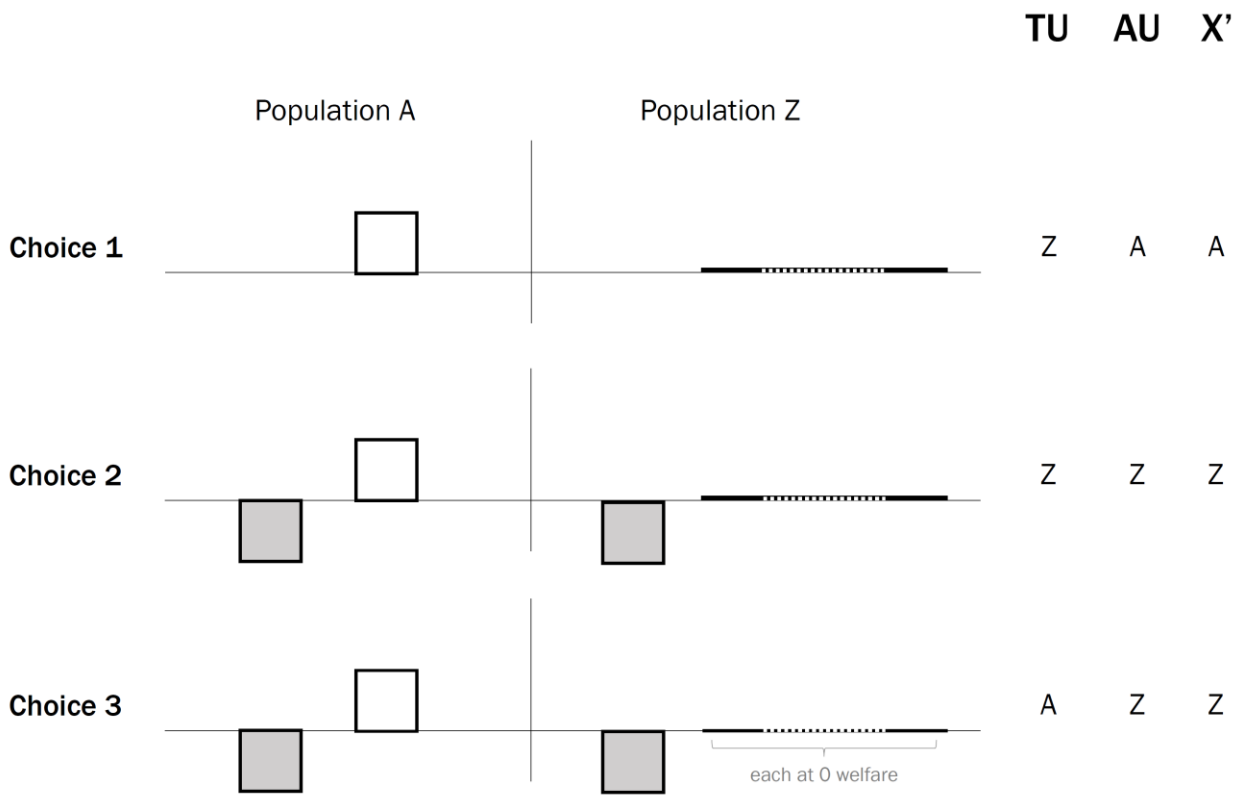


Figure 5: Different choices constitute instances of the Repugnant Conclusion for different axiologies, including not only Total Utilitarianism (TU), but also axiologies that are commonly assumed to avoid it such as Average Utilitarianism (AU) and Variable-Value Utilitarianism (X'). A choice is an instance of the Repugnant Conclusion for a given axiology if that axiology implies that Z is better than A (noted on the right side). Note that Choice 2 and Choice 3 involve additions to a base population (grey box) unaffected by the choice of which population to add. Choice 2 is like Choice 1, except that instead of a choice between the two distinct populations A and Z described in Figure 1, it is instead a choice between which of those same two populations A and Z to *add* to an existing base population. Choice 3 is like Choice 2, except that in Choice 3 the lives in Z each have wellbeing level of zero instead of barely positive.

In this figure, Choice 1 illustrates the canonical instance of the Repugnant Conclusion for Total Utilitarianism displayed in Figure 1 at the beginning of the paper, which is an implication of total utilitarianism (TU) but not of average utilitarianism (AU) nor

variable-value utilitarianism (X'). However, Choice 2 and Choice 3 are instances of the Repugnant Conclusion that are implied by both AU and X'. This shows that the Repugnant Conclusion is an implication of more axiologies than is commonly recognized. Furthermore, Choice 3 shows that there are instances of the Repugnant Conclusion that are avoided by TU, despite being implications of other axiologies such as AU and X' that are commonly assumed to avoid it. This shows that TU does not do worse with respect to the Repugnant Conclusion than AU or X', since it is not the case that TU implies all instances of the Repugnant conclusion, or that the instances implied by AU and X' are a proper subset of those implied by TU.[9] The population ethics literature has overlooked this phenomena, because it has restricted its attention to instances of the Repugnant Conclusion such as Choice 1 in which there is no base population to which a chosen population is added.

So, more axiologies imply the Repugnant Conclusion than is typically acknowledged in the literature, because the literature has mistakenly ignored the importance of instances of the Repugnant Conclusion with non-zero base population such as Choice 2 and Choice 3. This raises the question of how many more axiologies imply the Repugnant Conclusion once this mistake is corrected. The next step of the argument is to articulate a sufficient condition for being an instance of the Repugnant Conclusion that includes all the instances in Figure 5 and others like them that may or may not include a base population to which additions are made:

> **Sufficient Condition for the Repugnant Conclusion.** For any very high quality of life, any large number of high-quality lives, and any small positive level of well-being, there exists a number of low-positive-level lives and a (possibly empty) base population such that it is better to add the low-positive-level lives to the base population, rather than add the very-high-quality lives.

Budolfson and Spears 2018 prove that all Generalized Utilitarian Axiologies (and many more) satisfy the Sufficient Condition for the Repugnant Conclusion. In addition, they prove the stronger claim that these axiologies also satisfy:

---

[9] In noting the existence of multiple instances of the Repugnant Conclusion, each sharing the same fundamental repugnance, we follow Parfit (2016) who enumerates many heterogenous instances of the Repugnant Conclusion, writing of "another", "a", "this", and "another version of the" Repugnant Conclusion. Anglin (1977) notes the existence of examples like Choice 2 that show that "in some cases the average principle also leads to the Repugnant Conclusion" (p. 746).

Sufficient Condition for the Very Repugnant Conclusion. For any very high quality of life, any large number of high-quality lives, any very negative level of life full of suffering, and number (possibly zero) of negative lives, and any small positive level of well-being, there exists a number of low-positive-level lives and a (possibly empty) base population such that it is better to add the low-positive-level lives and the very-negative lives to the base population, rather than add the very-high-quality lives.

It is often claimed that avoiding the Very Repugnant Conclusion should be an even stronger desideratum than avoiding the Repugnant Conclusion, because the Very Repugnant Conclusion also involves preferring to add very negative lives (Arrhenius n.d.). The proof reveals that contrary to what has been assumed in the literature, all Generalized Utilitarian Axiologies (and many more) are also on equal footing with respect to even the Very Repugnant Conclusion.

The intuition behind the proof is that population axiologies imply the (Very) Repugnant Conclusion if they satisfy Convergence in Equivalence (stated in the previous section), together with a small number of other properties that all Generalized Utilitarian Axiologies assume are uncontroversial.[10] This is because satisfying Convergence in Equivalence means that an axiology *always cares about quantity to some non-zero extent*, whereas to avoid the RC an axiology has to deny this and insist that quantity effectively doesn't matter at least beyond some point. Convergence in Equivalence is satisfied by all Generalized Utilitarian Axiologies including total, average, and variable-value utilitarianism, and all prioritarian and egalitarian variations, as well as by many other axiologies. Thus, all of these axiologies (and more) imply instances of the (Very) Repugnant Conclusion.

---

[10] Budolfson and Spears 2018 show that any transitive, anonymous population axiology that satisfies Extended Egalitarian Dominance and Convergence in Equivalence implies the (Very) Repugnant Conclusion, as stated here. Critical-Level Generalized Utilitarianism does not satisfy these Extended Egalitarian Dominance unless it is standardized as in Broome (2004).

> **Extended Egalitarian Dominance**
> If population A is perfectly equal-in-wellbeing and is of greater size than population B, and every person in A has higher positive wellbeing than every person in B, then A is better than B (compare Arrhenius, n.d.).

They also define an "Extended Very Repugnant Conclusion", and show it is also implied by CLGU, maximin, and even more axiologies, including non-transitive, complete, and person-affecting axiologies, and show how even weaker conditions than Convergence in Signs can generate these implications.

Note that Convergence in Equivalence is not satisfied by rank-dependent views: maximin, for example, only cares about the worst-off person, while RDGU and Geometrism, in a way similar to maximin, care much more about some people than others. However, even rank-dependent axiologies imply an Extended Very Repugnant Conclusion, which fully retains (and arguably intensifies) the repugnant elements of Parfit's original statement of the Repugnant Conclusion. Non-Archimedean and Weak Value Superiority axiologies also do not satisfy Convergence in Equivalence. More generally, as Arrhenius (no date) demonstrates, such Non-Archimedean views have striking anti-egalitarian implications. They also imply their own versions of conclusions that are repugnant. For example, in the Non-Archimedean theory explored by Carlson in this volume, each person is described by an ordered pair of superior and inferior welfare. Although Carlson does not fully specify an aggregation method (he explores totaling within category), the essence of such a view is that it must satisfy within-category Pareto, privileging superior welfare (otherwise it would not avoid Parfit's Repugnant Conclusion). On any such view, in a same-number comparison, a population in which one person has some superior welfare (such as some opera and fine dining) and many people have high inferior welfare (such as plenty of muzak and potatoes), would be worse than a population in which the well-off person instead has slightly more superior welfare (such as one more minute of opera) and the rest of the population has lives full only of highly negative inferior welfare (such everyone's only experience, throughout their entire lives, is severe intestinal pain). This illustrates that any Non-Archimedean view will also have repugnant quantity-quality tradeoffs. Similar remarks apply to other welfarist axiologies, including rank-dependent axiologies (Budolfson and Spears 2018).

In a related vein, some authors have noted that the Repugnant Conclusion is fundamentally similar to same-number paradoxes of utilitarianism that welfarist consequentialists are prepared to accept (Cowen 1996, Ng 2013). Furthermore, even same-number axiologies (meaning social orderings that ignore the variable-number questions of population ethics) will inevitably have unintuitive implications when aggregating implications over unbounded spaces (Fleurbaey and Tungodden 2010). What these examples have in common is that, when aggregating and ranking over unbounded spaces, a factor that intuitively feels important can ultimately be outweighed by a factor that intuitively feels unimportant. Budolfson and Spears

extend the observations by Cowen, Ng, and Fleurbaey and Tungodden to explain why the Repugnant Conclusion, in one form or another, is implied by every generalized utilitarian axiology and more.

These examples suggest that what is fundamental to the sort of repugnance at issue in the Repugnant Conclusion may be the mere existence of a **Large Quantity-Quality Tradeoff**. This is surely one candidate for being the main intuition behind the counterintuitiveness of the Repugnant Conclusion. If this is the correct way to understand the nature of the Repugnant Conclusion – i.e. if Large Quantity-Quality Tradeoffs are sufficient for the Repugnant Conclusion – then it follows more or less directly that any plausible welfarist axiology implies instances of the Repugnant Conclusion, since any plausible welfarist axiology must presumably endorse the possibility of significant tradeoffs.

The upshot is that the Repugnant Conclusion does not tell against any axiology, because it cannot be avoided by any plausible axiology. This provides a new argument that the methodological requirement to avoid the Repugnant Conclusion should be dropped from population ethics. If so, this has implications for axiological theory, and also provides a further reason for thinking that the Repugnant Conclusion has no policy implications, at least within the familiar welfarist approach to policy evaluation.

6. Incorporating population ethics into ordinary policy analysis

For several decades, research in population ethics has focused on theoretical puzzles motivated in large part by the methodological goal of avoiding the Repugnant Conclusion. Recently, high-profile arguments have emerged that population ethics poses practical challenges to policy as well. As the discussion above indicates, we disagree: there is more convergence in the policy recommendations of competing population axiologies than is commonly understood. But what about cases in which there remains some amount of divergence?

We propose that even in these cases, population ethics can be incorporated into ordinary policy analysis without any difficulty. The method is straightforward: adopt a parameterized family of population axiologies, and investigate the robustness of policy choices to alternative parameter choices. This is perhaps the most common method in

actual policy analysis of representing normative uncertainty, is already routinely done for other dimensions of normative uncertainty. For example, climate policy models such as William Nordhaus' DICE and RICE models choose policy in order to maximize a (total) utilitarian policy objective (called a social welfare function in this literature) (Nordhaus 2017, Nordhaus 2019). This objective is broadly utilitarian, but depends upon two parameters: η, which parameterizes the degree of inequality aversion (and thus the connection between consumption and wellbeing; mathematically, the degree of concavity of diminishing marginal wellbeing as a function of consumption), and ρ, which parameterizes the rate of pure time preference, according to which future wellbeing is discounted simply because it is in the future. Climate policy analysis routinely investigates whether recommendations hold for a *range* of values of these parameters, meaning that policy can be chosen without resolving normative uncertainty about utilitarianism and prioritarianism, what the correct intertemporal savings principle is, whether there are agent-relative permissions to give less weight to people in the far future, and so on. Similarly, health policy decisions are often investigated over a range of quantities for the value of a statistical life, or a disability-adjusted life year. Education and taxation policies are chosen that work for a variety of elasticities of taxable income. In all these cases, normative uncertainty is acknowledged and explicitly investigated in policy analysis, by parameterizing the key normative dimensions of the analysis and testing the sensitivity of optimal policy to the range of reasonable uncertainty about those dimensions.

It is straightforward to extend this method to population ethics. Consider this broad family of population axiologies: $n^{\alpha}[\bar{u} - c]$, where n is the population size and $\bar{u}$ is the equally-distributed-equivalent wellbeing[11] (which could incorporate utilitarianism, prioritarianism, or egalitarianism). In a policy analysis, population ethics could be roughly parameterized as a critical level $c \geq 0$, and as the exponent $\alpha \in [0,1]$, where $\alpha = 0$ is averagism, $\alpha = 1$ is totalism, and values in between are variable-value forms of each of these. Many policy recommendations — such as complete decarbonization by 2150 — will be uniformly recommended by all plausible values of α and c. As a result, policy can be chosen that is robust to a wide range of future theoretical resolutions of the puzzles of the Repugnant Conclusion and population ethics, more broadly. In

[11] For simplicity, the critical level is not included within the equally-distributed-equivalent calculation, but without loss of generality this could be accommodated in this framework.

other cases, we could discover more divergence. But then by implementing this method we can learn the magnitude of that divergence and rigorously inform our decision making under normative uncertainty, in the same way we inform decisions under other kinds of normative uncertainty in ordinary policy analysis. Either way, uncertainty about population ethics does not provide an important barrier to policymaking.[12]

[12] We acknowledge with gratitude helpful comments from Gustaf Arrhenius, Krister Bykvist, Tim Campbell, Tom Christiano, Diane Coffey, Mike Geruso, Johan Gustafsson, Vasudha Jain, Jonathan Riley, Melinda Roberts, Orri Stefánsson, Chad Van Schoelandt, Sangita Vyas, Andrew Williams and participants in seminars at IFFS, ANU, the Paris School of Economics, and the 2019 PPE conference in New Orleans.

References

Anglin, Bill, 1977. The repugnant conclusion. *Canadian Journal of Philosophy*, 7(4), pp.745-754.

Arrhenius, Gustaf (n.d.) *Population Ethics: The Challenge of Future Generations*: unpublished typescript (July 2017).

Arrhenius, Gustaf and H. Orri Stefánsson (2018) "Population ethics under risk," working paper, Institute for Futures Studies.

Arrhenius, Gustaf, Mark Budolfson, and Dean Spears. 2019. Does Climate Change Policy Depend Importantly on Population Ethics? Deflationary Responses to the Challenges of Population Ethics for Public Policy, in Budolfson, M., McPherson, T., and Plunkett, D. eds. *Climate Change and Philosophy*. Oxford UP.

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

Babiarz, K.S., Ma, P., Miller, G. and Song, S., 2018. The Limits (and Human Costs) of Population Policy: Fertility Decline and Sex Selection in China under Mao (No. w25130). National Bureau of Economic Research.

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

Blackorby, Charles, Walter Bossert, and David J Donaldson (2005) *Population issues in social choice theory, welfare economics, and ethics*, No. 39: Cambridge University Press.

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

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

Broome, John (2004) *Weighing lives*, Oxford UP.

Broome, John. 2012. *Climate Matters: Ethics in a Warming World*. Norton.

Budolfson, Mark. 2018. Market Failure, the Tragedy of the Commons, and Default Libertarianism in Contemporary Economics and Policy, in the *Oxford Handbook of Freedom*. Oxford UP.

Budolfson, Mark and Dean Spears. 2019a. Population momentum, population ethics, and the prospects for fertility policy as climate mitigation policy. Working paper: UVM and UT-Austin.

Budolfson, Mark and Dean Spears. 2019b. Methods for quantifying animal wellbeing and estimating optimal tradeoffs against human wellbeing — and lessons for axiology, including new arguments for separability. Working paper: UVM and UT-Austin.

Budolfson, Mark and Dean Spears. 2018. Why the Repugnant Conclusion is Inescapable. Working paper: Princeton University Climate Futures Initiative.

Cai, Y., 2010. China's below-replacement fertility: government policy or socioeconomic development?. *Population and Development Review*, *36*(3), pp.419-440.

Coffey, D. and Spears, D. 2017. *Where India Goes: Abandoned Toilets, Stunted Development, and the Costs of Caste.* Harper Collins.

Connelly, M.J., 2009. *Fatal misconception: The struggle to control world population*. Harvard University Press.

Cowen, T. (1996) "What do we Learn from the Repugnant Conclusion?" *Ethics*, Vol. 106, pp.754–775.

Deaton, A. 2013. *The Great Escape.* Princeton UP.

Fleurbaey, M., 2010. Assessing risky social situations. *Journal of Political Economy*, *118*(4), pp.649-680.

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

Gerland, P., Raftery, A.E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B.K., Chunn, J., Lalic, N. and Bay, G., 2014. World population stabilization unlikely this century. *Science*, *346*(6206), pp.234-237.

Geruso, M. and Spears, D. 2018. “Heat, Humidity, and Infant Mortality in the Developing World,” IZA Discussion Paper 11717.

Goodkind, D., 2017. The astonishing population averted by China’s birth restrictions: Estimates, nightmares, and reprogrammed ambitions. *Demography*, *54*(4), pp.1375-1400.

Greaves, H. 2017 “Population axiology,” *Philosophy Compass*, Vol. 12.

Greenhalgh, S., 2008. *Just one child: Science and policy in Deng’s China*. Univ of California Press.

Huemer, M. 2008. ‘‘In Defence of Repugnance,’’ *Mind*, 117 (468): 899–933

Lam, D. 2011. “How the world survived the population bomb: Lessons from 50 years of extraordinary demographic history.” *Demography*, *48*(4), pp.1231-1262.

Nebel, J. 2019. “An Intrapersonal Addition Paradox.” *Ethics, 129*(2), pp. 309-343.

Ng, Y.K. 1989. “What Should We Do About Future Generations?: Impossibility of Parfit's Theory X.” *Economics & Philosophy*, *5*(2), pp.235-253.

Ng, Y.K. 2013. “Number dampened utilitarianism”, in Crimmins, J. ed. The *Bloomsbury Encyclopedia of Utilitarianism*, Bloomsbury.

Nordhaus, W.D., 2017. Revisiting the social cost of carbon. *Proceedings of the National Academy of Sciences*, *114*(7), pp.1518-1523.

Nordhaus, W. 2019. Climate Change: The Ultimate Challenge for Economics, Nobel Prize lecture.

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

Parfit, Derek., 2016. Can we avoid the repugnant conclusion?. *Theoria*, 82(2), pp.110-127.

Pritchett, Lant, 1994. Desired Fertility and the Impact of Population Policies. *Population and Development Review*, pp.1-55.

Pummer, T., 2013, “Intuitions about large number cases” *Analysis*, 73 (1): 80–98.

Roberts, Melinda A. (2015) “Population axiology,” in the *Oxford Handbook of Value Theory*, pp.399–423.

Rosling, Hans. 2018. *Factfulness.* Flatiron.

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*, Vol.114, pp. 12338–12343.

Sider, Ted (1991) “Might theory X be a theory of diminishing marginal value?” *Analysis*, Vol. 51, pp. 265–271.

Spears, Dean (2018) “Probabilistic future people: Social criteria for evaluating population prospects,” working paper, UT-Austin.

Spears, Dean and Mark Budolfson. 2019a. Radical Agreement in Population Ethics—and Implications for Social Policy Analysis. Working paper: UVM and UT-Austin.

Spears, Dean and Mark Budolfson. 2019b. The repugnant conclusion cannot be escaped: Proofs and implications. Working paper: UVM and UT-Austin.

Tannsjo, T., 2002. Why we ought to accept the repugnant conclusion. *Utilitas*, *14*(3), pp.339-359.

Zhao, Z. and Zhang, G., 2018. Socioeconomic Factors Have Been the Major Driving Force of China’s Fertility Changes Since the Mid-1990s. *Demography*, 55(2), pp.733-742.