



Population ethics under risk

Gustaf Arrhenius^{1,2} · H. Orri Stefánsson^{1,2,3}

Received: 8 July 2020 / Accepted: 21 January 2023 / Published online: 18 March 2023
© The Author(s) 2023

Abstract

Population axiology concerns how to rank populations by the relation “is socially preferred to”. So far, population ethicists have (with important exceptions) focused less on the question of how to rank population prospects, that is, alternatives that contain uncertainty as to which population they will bring about. Most public policy choices, however, are decisions under uncertainty, including policy choices that affect the size of a population (such as climate policy choices). Here, we shall address the question of how to rank population prospects by the relation “is socially preferred to”. We start by illustrating how well-known population axiologies can be extended to population prospect axiologies. And we show that new problems arise when extending population axiologies to prospects. In particular, traditional population axiologies lead to prospect-versions of the problems that they are praised for avoiding in the risk-free settings. Moreover, we show how the axiom of State-Wise Dominance allow us to extend any impossibility theorem in population axiology to impossibility theorems for *non-trivial* population prospects, that is, prospects that confer probabilities *strictly between* zero and one on different populations. Finally, we formulate impossibility results that only involve probabilistic axioms.

We would like to thank Matt Adler, Geir Asheim, Krister Bykvist, Tim Campbell, Marc Fleurbaey, Johan Gustafsson, Robert Goodin, Julia Mosquera, Jake Nebel, Wlodek Rabinowicz, Daniel Ramöller, Dean Spears, Alex Voorhoeve, Stéphane Zuber, and three referees for this journal for very helpful comments and suggestions. Thanks also to the audiences at the New Scholarship on Population Ethics, Duke University, April 24, 2017; Sustainability and Population Ethics, Iméra Aix Marseille Université, 3 July, 2017; the workshop Population Ethics, Institute for Futures Study, May 4, 2017; Boulder Formal Value Theory Workshop, March 19, 2018; and workshop on Population, Social Welfare, and Climate Change, at UT Austin, February 7 2020. Financial support from Riksbankens Jubileumsfond (through grant number M17-0372:1 for Arrhenius and a Pro Futura Scientia XIII grant for Stefánsson) is gratefully acknowledged.

✉ H. Orri Stefánsson
orri.stefansson@philosophy.su.se

Gustaf Arrhenius
gustaf.arrhenius@iffs.se

¹ Institute for Futures Studies, Stockholm, Sweden

² Department of Philosophy, Stockholm University, Stockholm, Sweden

³ Swedish Collegium for Advanced Study, Uppsala, Sweden

1 Introduction

Population axiology concerns how to evaluate populations in terms of their “moral goodness” as philosophers call it, that is, how to rank populations by the relation “is socially preferred to”, to use terminology more familiar in social choice theory and welfare economics. So far, discussions of population axiology has focused less on the question of how to rank alternatives that contain uncertainty as to which population they will bring about.¹ The adequacy conditions (i.e., “axioms”) that have been discussed mostly pertain to the ranking of risk-free outcomes and are silent in regard to how to rank (non-trivial) risky prospects. In other words, questions about what we call *population prospect axiology*—that is, questions about how to rank what we call *population prospects*—have, in our view, been underexplored.

This exclusion of uncertainty is unfortunate since most public policies that could affect the number and identities of future people contain a great deal of uncertainty as to which population they will bring about.² For instance, as the Intergovernmental Panel on Climate Change acknowledged in their Fifth Assessment Report,³ we should expect that the climate policies we choose will affect the size of the world population. But obviously, climate policies contain a great deal of uncertainty, for instance, uncertainty about what type of population they will bring about.⁴

Since the birth of modern population axiology, starting with Parfit’s informal paradoxes summarised in Parfit (1984), researchers in this area have uncovered a number of *impossibility results* that establish that no population axiology can consistently satisfy a small number of conditions that many find compelling as adequacy conditions on any such axiology.⁵ Unfortunately, it turns out that the situation is even worse when it comes to population *prospect axiology*. As we show below, new problems appear when one tries to turn a population axiology into a population prospect axiology. Perhaps most interestingly, we shall see that natural extensions of traditional (risk-free) population axiologies lead to prospect-versions of the problems that they are praised for avoiding in the risk-free setting.

Moreover, we show that an instance of the familiar axiom of State-Wise Dominance opens the door to impossibility results for population prospect axiology that are analogous to the traditional impossibility results for (risk-free) population axiology. Importantly, these results cannot be derived just from the (risk-free) adequacy conditions involved in the traditional impossibility theorems, and they hold for

¹ For important exceptions, see Hammond (1988), Blackorby et al. (1998), Blackorby et al. (2005), Roberts (2007), Asheim and Zuber (2016), Nebel (2019), Budolfson and Spears (2018), Spears and Budolfson (2019), McCarthy et al. (2020), Thornley (2021). Voorhoeve and Fleurbaey (2016) and Nebel (2017) discuss prospects involving populations with different people but the same number of people.

² Contrary to the convention in economics, we use “uncertainty” and “risk” interchangeably throughout this paper. So, a situation of risk, and a situation of uncertainty, is one where one knows what outcome one’s choices *might* result in, but does not know which outcome one’s choice *will* result in.

³ Available at <https://www.ipcc.ch/assessment-report/ar5/>.

⁴ For further examples, see Broome (1992, 2004, 2010, 2015). See also Arrhenius (forthcoming, ch. 1) for a discussion.

⁵ See e.g., Ng (1989), Arrhenius (2000a, 2000b, 2011), Blackorby et al. (2005).

non-trivial prospects (as well as for trivial ones), that is, for prospects that confer probabilities *strictly between* zero and one on different populations. So, with this adequacy condition on prospects, all the well-known problems from risk-free population axiology are reproduced in the risky setting; and, in addition, new problems appear when risk is introduced.

Now, we know that there are some multi-person decision problems where a solution can be found only if we introduce probabilities. In particular, there are some well-known problems in game theory where a *Nash equilibrium* exists only if the “players” are allowed to randomise between their available “pure” (i.e., non-probabilistic) strategies (von Neumann and Morgenstern 1944). In a similar vein, one might have hoped that a solution could be found to the problems of population ethics by introducing probabilities. This hope might seem further supported by the fact that the traditional adequacy conditions for population axiology say nothing about how to rank what we above called “non-trivial” prospects, that is, prospects that confer probabilities strictly between zero and one on different populations (as we further explain below).

An axiology for non-trivial population prospects that avoids the traditional risk-free impossibility results would indeed be a great victory from a practical point of view. Policies that could affect the size of a population typically contain considerable uncertainty, as previously mentioned. In fact, it is hard to think of any policy (population-affecting or not) that is certain to deliver a particular outcome. Therefore, one might conjecture that if we have an adequate population prospect axiology, then we will be able to make consistent population policy choices, notwithstanding the traditional impossibility results for risk-free populations. But as we shall soon see, the traditional impossibility results more-or-less reappear when risk is introduced, which, as previously mentioned, moreover brings with it some additional problems.

2 Framework and Background Assumptions

A natural starting point in the search for a population prospect axiology is to consider some extensions of standard population axiologies to situations of uncertainty. But first, we need to introduce a few terms and distinctions used in traditional, risk-free, population axiology, which we will use when formulating the aforementioned extensions.

In traditional population axiology, one often distinguishes between the welfare of a person and the contributive value of a person’s welfare (or well-being; we shall use these two terms interchangeably). By the contributive value of a person’s welfare we shall mean the value that a person’s welfare contributes to the value of a population of which it is a member. More exactly, the contributive value of a person i ’s welfare relative to a population \mathbf{x} , of which i is a member, is the difference in value between \mathbf{x} and the population consisting of all the \mathbf{x} -people except i . This distinction is useful for population prospect axiology too, for we can now reformulate one of our main questions as “What is the contributive value of a possible person’s welfare to a prospect?”

Now to our formal framework, which follows closely that of Asheim and Zuber (2014), although suitably extended to prospects based on the framework of Savage (1954). Let \mathbb{N} denote the set of natural numbers and \mathbb{R} the set of real numbers. $\mathbb{R}^- \subset \mathbb{R}$ and $\mathbb{R}^+ \subset \mathbb{R}$ are the sets of negative and positive real numbers. Let $\mathbf{X} = \bigcup_{n \in \mathbb{N}} \mathbb{R}^n$ denote the set of possible finite distributions of lifetime well-being. More formally, \mathbf{X} is a set of vectors of real numbers, where each number (“utility”) represents the lifetime well-being of some person. A generic such vector for a population of m people is denoted $\mathbf{x} = (x_1, \dots, x_m)$ or $\mathbf{y} = (y_1, \dots, y_n)$, where, say, x_i denotes the lifetime well-being of individual i . Informally, we shall talk of a vector like this as a *population*. $\mathbf{X}^- \subset \mathbf{X}$ and $\mathbf{X}^+ \subset \mathbf{X}$ are the subsets of such vectors where *every* number in the vector is negative respectively positive. The size of population \mathbf{x} is denoted by $n(\mathbf{x})$ (and will, as mentioned, always be finite). For any vector \mathbf{x} , we write the average value of its elements as \bar{x} . So, \bar{x} should be interpreted as the average lifetime welfare of people in population \mathbf{x} .

Let $(z)_n \in \mathbb{R}^n$ denote the perfectly equal population where all n individuals have lifetime well-being z . We let (\mathbf{x}, \mathbf{y}) denote the population that results when population $\mathbf{x} \in \mathbf{X}$ is conjoined with population $\mathbf{y} \in \mathbf{X}$. And let $(\mathbf{x}, (z)_n)$ denote population $\mathbf{x} \in \mathbf{X}$ with n added individuals that all have lifetime well-being z .

The well-being measure is normalised around the “neutral” life. More formally, if i 's life in \mathbf{x} is neutral, then $x_i = 0$, while if i 's life in \mathbf{x} is better for i than a neutral life, then i has positive welfare in \mathbf{x} , i.e., $x_i > 0$, but if i 's life in \mathbf{x} is worse for i than a neutral life, then i has negative welfare in \mathbf{x} , i.e., $x_i < 0$. We shall remain agnostic about what constitutes well-being and what determines whether a life is neutral (but for different suggestions for the latter, see e.g. Arrhenius, [forthcoming](#), ch. 2 and 9, Arrhenius 2000a, Broome, 2004, and Parfit, 1984, pp. 357–358, and appendix G).

To simplify the formal presentation of views and conditions, we shall assume the standard *anonymity* assumption, which holds that the “social welfare relations” that we study are invariant under permutations of the vectors in \mathbf{X} . So, for instance, let \mathbf{x}' be the vector that results when the lifetime well-being of i and j in \mathbf{x} are switched. Then the social welfare relations that we shall consider are all indifferent between \mathbf{x} and \mathbf{x}' , that is, they deem these two populations to be equally good. Intuitively, this means that it does not matter *who* receives what welfare; all that matters is how welfare quantities are distributed. This assumption rules out some views in the literature that are called “person affecting” (cf. Glover, 1977, Arrhenius, 2009, Roberts, 2010) as well as desert-based views (Arrhenius 2003a; Feldman 1997). But we should emphasise that anonymity is just a simplifying assumption. Just like the original results in the risk-free setting, the results of this paper can be established without an anonymity assumption, but the argument would then involve more complicated formulations.

We let \lesssim on \mathbf{X} denote a “social welfare relation” on \mathbf{X} , such that for any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, $\mathbf{x} \lesssim \mathbf{y}$ means that \mathbf{y} is at least as socially preferable as \mathbf{x} , or, as we shall often put it, \mathbf{y} is at least as *good* as \mathbf{x} . (So, \lesssim is what some philosophers would call a “moral goodness” relation, but we shall from now on refer to it as a social welfare relation.) To simplify the discussion we shall typically assume that the social welfare relation is transitive, reflexive, and complete, which means that the relation generates a social welfare *order*. It is worth noting, however, that for the impossibility results

that we present in section 4 a *quasi-order* suffices, that is, completeness need not be assumed. In other words, these results are consistent with some populations being incommensurable in value.⁶ The strict relation, \prec , and indifference, \sim , are respectively the asymmetric and symmetric counterparts of \preceq . When we speak of a “social planner”, we mean the (possibly group) agent whose preference is represented by \preceq .

Although we shall typically assume that the relevant probabilities are “objective”, or given to the social planner, as in the von Neumann and Morgenstern (1944) framework, we use a more general Savage (1954)-inspired framework—which does not assume the probabilities to be given to the decision-maker—to turn the traditional population axiologies into population prospect axiologies. To that end, let \mathbf{S} be a set of (non-overlapping) *states of the world*, that is, resolutions of uncertainty, with generic elements s_i, s_j . For simplicity, we assume \mathbf{S} to be finite. Let \mathbf{P} be a set of functions from \mathbf{S} into \mathbf{X} . Generic elements of \mathbf{P} will be denoted P, Q , or P_1, P_2 , etc. We interpret the elements in \mathbf{P} as *population prospects*. $P(s_i) \in \mathbf{X}$ is the well-being vector, that is, the population, that results from prospect P if s_i is the actual state of the world. Finally, we assume that the social planner has a probability measure over \mathbf{S} , that is, there is a function p from \mathbf{S} into $[0,1]$ such that $\sum_{i=1} p(s_i) = 1$, which quantifies the social planner’s uncertainty (which may but need not correspond to some objective probability distribution). We use \preceq for weak social welfare relation over \mathbf{P} , with the corresponding strict relation, \prec , and indifference, \sim .

It will be useful to work with *events*, that is, unions of states, when stating some conditions. To that end, let \mathbf{E} be a Boolean algebra of events generated from \mathbf{S} , that is, the set of all possible sets of states in \mathbf{S} , closed under the Boolean operators. Generic elements of \mathbf{E} will be denoted $E, \neg E, F$, etc. Since we assume the social planner to be probabilistically coherent, her probability function extends straightforwardly from \mathbf{S} to \mathbf{E} . Moreover, probabilistic coherence ensures that if $E \subset \mathbf{E}$ then $p(E) < 1$ and $p(\neg E) > 0$, which will be convenient when stating some of the below conditions. Finally, to further simplify notation, when we for instance write $P(E) = \mathbf{x}$, this means that *every* state in E results in \mathbf{x} given P , that is, for any $s_i \in E$, $P(s_i) = \mathbf{x}$.

Later we shall focus much of our attention on what we call *non-trivial* prospects. Informally, a prospect is non-trivial if it confers a positive probability on at least two distinct populations; otherwise it is said to be trivial. More formally:

Definition (*Non-trivial prospect*) A prospect, $P \in \mathbf{P}$ is called *non-trivial* just in case there are $P(s_i), P(s_j) \in \mathbf{X}$ where $P(s_i) \neq P(s_j)$ and $p(s_i) > 0, p(s_j) > 0$.

We let $\mathbf{P}_{nt} \subset \mathbf{P}$ denote the set of non-trivial population prospects.

⁶ For recent discussions of limited, or “imprecise”, comparability in the context of population ethics, see Arrhenius (2016) and Parfit (2016). See also Blackorby et al. (2005) on incomparability between populations.

3 Population Prospect Axiologies

Let us now consider how the traditional population axiologies could be extended to situations of uncertainty, using the framework and terminology from the last section.⁷ We start with Total Utilitarianism, according to which the contributive value of a person's welfare equals her welfare,⁸ and the value of a population is calculated by summing the welfare of all people in the population.

Total Utilitarianism (TU) For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$:

$$\mathbf{x} \lesssim \mathbf{y} \Leftrightarrow TU(\mathbf{x}) = \sum_i x_i \leq TU(\mathbf{y}) = \sum_i y_i$$

A natural way to turn Total Utilitarianism into a population prospect axiology is to sum up weighted Total Utilitarian values of all the possible populations that a prospect might result in, where the weight on the value of each population is determined by the probability that the prospect will result in that population.⁹ Let us call the resulting axiology *Prospect Total Utilitarianism*.

Prospect Total Utilitarianism (PTU) For any $P, Q \in \mathbf{P}$:

$$P \lesssim Q \Leftrightarrow \sum_i p(s_i)TU(P(s_i)) \leq \sum_i p(s_i)TU(Q(s_i))$$

One of the great challenges that examinations of different size populations have raised for Total Utilitarianism, is how to respond to the fact that the latter gives rise to what Parfit (1984) called the *Repugnant Conclusion*. Informally, this conclusion

⁷ To focus the discussion—and due to limited space—we shall set aside rank-dependent views, and other views that have been called “variable-value views”, early version of which were discussed by Hurka (1983), Ng (1989), Sider (1991). More recent discussions of such views include Zuber and Asheim (2012), Asheim and Zuber (2014, 2016), and Pivato (2020).

⁸ Strictly speaking, the contributive value of a person's welfare can be a linear transformation of her welfare, according to Total Utilitarianism; but it can neither be a concave nor a convex transformation of her welfare.

⁹ The population prospects axiologies we formulate treat risk and uncertainty as standard Expected Utility Theory does. There are reasons for being sceptical of that treatment; for instance, since it does not allow for risk aversion with respect to utility. However, to keep things simple, we shall assume the standard expected utility treatment of risk. It is easy to see that the results we shall discuss would still hold if we assumed some of the minimally normative alternative treatments of risk that have recently been suggested, that is, treatments that satisfy transitivity and stochastic dominance, such as the theories suggested by Quiggin (1982), Tversky and Kahneman (1992), Buchak (2013) and Stefánsson and Bradley (2015, 2019). In contrast, our results do not hold if one uses, say, the original Prospect Theory (Kahneman and Tversky 1979), which cannot simultaneously satisfy stochastic dominance and transitivity, or Regret Theory (Bell 1982; Loomes and Sugden 1982), which only satisfies transitivity w.r.t. a triple of alternatives if at least one alternative in the triple strictly state-wise dominates another alternative in the triple (Diecidue and Somasundaram 2017). While the latter theories may be accurate descriptions of how people actually rank prospects, few if any scholars would suggest them as theories about how one ideally should rank prospects. Therefore, we think that we can justifiably set them aside for the present purposes, where the aim is to discuss how one *should* rank population prospects.

is that for any perfectly equal population (no matter how well off are its members), there is a better population consisting of people with very low positive welfare. More formally¹⁰:

Repugnant Conclusion For all $y, z \in \mathbb{R}$, where $y > z > 0$, and for any $k \in \mathbb{N}$, there is an $n \in \mathbb{N}$, where $n > k$, and such that $(y)_k < (z)_n$.

It should be evident that the prospect version of Total Utilitarianism also leads to the Repugnant Conclusion not only in non-risky situations but also in risky ones. For instance, for any prospect P , all of whose possible outcomes are populations where people enjoy very high welfare, there exists some prospect Q , all of whose possible outcomes are larger populations where people have very low positive welfare, such that Prospect Total Utilitarianism will recommend choosing the latter. The result, if we follow the recommendation of Prospect Total Utilitarianism, is that we end up with a large population where people have very low positive welfare when we had the option of some smaller population where people enjoy very high welfare—an instance of the Repugnant Conclusion. Likewise for other potentially counterintuitive implications of Total Utilitarianism in non-risky settings, such as the Very Repugnant Conclusion.¹¹

Moreover, Prospect Total Utilitarianism violates probabilistic analogues of the traditional (risk-free) Quality condition which essentially rules out the Repugnant Conclusion (and which we state in section 4). Let $\mathbb{H} \subset \mathbb{R}^+$ be welfare levels that are very high—such that a life at a level in \mathbb{H} would be deemed excellent—while $\mathbb{L} \subset \mathbb{R}^+$ are welfare levels that are very low but positive—such that a life at a level in \mathbb{L} would be deemed barely worth living. Consider the following probabilistic analogue of the traditional Quality condition, which can be derived from the non-risky condition with help of the traditional State-Wise Dominance principle from decision theory that we discuss in section 4.¹² Informally, what we shall call *Probabilistic Quality* says that there is some very high welfare level z and some number n of people at that level, such that for any very low but positive welfare level y and any number m of people at that level, if prospects P and Q result in the same population in some state(s), but in states where they result in different populations, P results in a population where n people are at the very high level z whereas Q results in a

¹⁰ It is worth noting that this formulation is more general than Parfit's (1984): "For any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better even though its members have lives that are barely worth living" (388). Since the reason we discuss the Repugnant Conclusion in this article is only to illustrate what some have taken to be an undesirable implication of Total Utilitarianism, no harm is done by focusing on the more general version, which Total Utilitarianism of course implies.

¹¹ Informally, the Very Repugnant Conclusion states that for any perfectly equal high-welfare population, and for any number of people with very negative welfare, there is a population consisting of the people with negative welfare and people with very low positive welfare which is better than the high welfare population (Arrhenius 2003b).

¹² Recall that $P(E) = \mathbf{x}$ means for any $s_i \in E$, $P(s_i) = \mathbf{x}$; hence, we can apply State-Wise Dominance when, say, $P(E) = \mathbf{x}$ and $P(\neg E) = (z)_n$.

population where m people are at the very low level y , then P is at least as good as Q . More formally:

Probabilistic Quality There is a $z \in \mathbb{H}$ and an $n \in \mathbb{N}$, such that for any $\mathbf{x} \in \mathbf{X}$, any $E \subseteq \mathbf{E}$, any $y \in \mathbb{L}$ and any $m \in \mathbb{N}$, if $Q(E) = \mathbf{x}$ and $Q(\neg E) = (z)_n$, while $P(E) = \mathbf{x}$ and $P(\neg E) = (y)_m$, then $P \lesssim Q$.

Prospect Total Utilitarianism of course violates Probabilistic Quality. The expected utility framework we have assumed when extending Total Utilitarianism to Prospect Total Utilitarianism implies that the probability and utility of \mathbf{x} can be ignored when comparing P with Q (this follows from the aforementioned dominance condition). Hence, since m can be very large, compared with n , Prospect Total Utilitarianism violates Probabilistic Quality, for the same reason that Total Utilitarianism leads to the Repugnant Conclusion.¹³

In comparison with other axiologies, Total Utilitarianism arguably handles (non-risky) cases involving negative welfare better than most alternative axiologies, such as Average and Critical-Level Utilitarianism (more on which below). For example, Total Utilitarianism clearly avoids the Very Sadistic Conclusion which is implied by Critical-Level Utilitarianism.¹⁴ Informally, the Very Sadistic Conclusion is that for any population with negative welfare, there is a worse population with positive welfare. More formally:

Very Sadistic Conclusion For any $\mathbf{x} \in \mathbf{X}^-$ there is a $\mathbf{y} \in \mathbf{X}^+$ such that $\mathbf{y} < \mathbf{x}$.

While Total Utilitarianism avoids the Very Sadistic Conclusion, Prospect Total Utilitarianism leads to what we call the *Risky Very Sadistic Conclusion*.

Risky Very Sadistic Conclusion For any $E \subset \mathbf{E}$, for any $\mathbf{x} \in \mathbf{X}^-$, for any $\mathbf{y} \in \mathbf{X}^+$, and for any $\gamma \in \mathbb{R}^+$, there is a $z \in \mathbb{R}$ such that $z < \gamma$ and an $m \in \mathbb{N}$ such that if $P(E) = \mathbf{x}$ and $P(\neg E) = (z)_m$, while the *trivial* prospect Q for sure results in \mathbf{y} , then $Q \prec P$.

Here is why Prospect Total Utilitarianism implies the Risky Very Sadistic Conclusion. For any probability $p(E) < 1$, however high, if m is sufficiently great, then the expected total welfare of prospect P , which has probability p of resulting in \mathbf{x} and probability $1 - p$ of resulting in $(z)_m$, will be higher than the expected total welfare of prospect Q , which results in \mathbf{y} for sure. Thus, Prospect Total Utilitarianism deems P better than Q . This holds irrespective of how much the \mathbf{x} -people suffer and of how many they are. In other words, while Total Utilitarianism avoids the Very Sadistic Conclusion, Prospect Total Utilitarianism recommends prospects that will almost certainly result in a population with negative welfare even though a population with

¹³ We note that Budolfson and Spears (2018) discuss a similar result, involving a probabilistic version of the Very Repugnant Conclusion.

¹⁴ For a discussion, see Arrhenius (2000a, 2000b, forthcoming). Average Utilitarianism has similar problems.

positive welfare was available (hence the name of the conclusion). Below we prove a general theorem of which this result is a special case.¹⁵

Now, there is of course a sense in which the above conclusion (as well as the aforementioned general result) is just an instance of a general implication of Expected Utility Theory (EUT)—and, in particular, its probabilistic continuity axiom—that has nothing specifically to do with population ethics.¹⁶ We however make two points about this observation. First, we think that while one might see the Risky Very Sadistic Conclusion as a special case of a more general implication of EUT, that does not change the fact that the special case in question is illuminating for debates in population ethics. In particular, it teaches us something important about Total Utilitarianism, namely, that while the theory is praised for avoiding the Very Sadistic Conclusion, the most natural prospect-extension of the theory comes arbitrarily close to the very same conclusion, in the sense of recommending, for any population with negative welfare, a choice that has an arbitrarily high probability of resulting in that population, even though a population with positive welfare was available. So, the theory comes arbitrarily close to recommending the Very Sadistic Conclusion.¹⁷

Second, it is worth noting that two of the most influential arguments in favour of Utilitarianism assume continuity of the kind that leads total utilitarians to the Risky Very Sadistic Conclusion; this is true both of arguments for Average Utilitarianism (Harsanyi 1953), and Total Utilitarianism (Harsanyi 1955; Hammond 1987, 1988).¹⁸ So, although continuity in probabilities, as assumed by Expected Utility Theory, is largely to blame for the Risky Very Sadistic Conclusion—as well as the Risky Repugnant Conclusion, which we discuss below—utilitarians (in particular Prospect Utilitarians) presumably will be wary of responding by denying such continuity.¹⁹

Faced with the Repugnant Conclusion and other similar problems, a Utilitarian might be tempted to opt for a non-total version of the theory, such as Average Utilitarianism or Critical-Level Utilitarianism, which can respectively be extended to Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism. It is easy to see that these theories avoid the Repugnant Conclusion. However, as is

¹⁵ Some may dislike the name of this conclusion for the following reason: Total utilitarians think that the population with negative welfare is the worst of the possible populations involved in the above prospect; hence, the Risky Very Sadistic Conclusion differs in an important respect from the ordinary Very Sadistic Conclusion, where critical-level utilitarians think that a population with negative welfare is *better* than one with positive welfare. Nevertheless, we think that the label is apt for the reason explained above: The Very Risky Sadistic Conclusion will almost certainly result in an outcome analogous to the outcome of the Very Sadistic Conclusion.

¹⁶ We thank Johan Gustafsson for pressing us on this point.

¹⁷ We also note that this implication of Prospect Total Utilitarianism, as well as the implication of Prospect Average Utilitarianism that we discuss below, supports the suggestions by Spears and Budolfson (2019) and Budolfson and Spears (2018) that both “repugnance” and “sadism”, given their general understanding of these terms, is harder to avoid than what many scholars seem to think (as also pointed out by Arrhenius 2000a, forthcoming).

¹⁸ Note however that strictly speaking, these are only arguments for Utilitarianism in a fixed number setting.

¹⁹ On the other hand, we should also acknowledge that continuity in probabilities is not *essential* to axiomatic Utilitarianism, as observed by McCarthy et al. (2020).

well known, these alternatives to Total Utilitarianism have, in the risk-free setting, their own counterintuitive conclusions to deal with; conclusions that are arguably at least as counterintuitive as the Repugnant Conclusion.²⁰ It should be evident that since Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism are extensions of Average Utilitarianism and Critical Level Utilitarianism to risky situations, they inherit the problems of their ancestors for *trivial* prospects.²¹ Moreover, it turns out that Prospect Average Utilitarianism and Prospect Critical Level Utilitarianism—and, in fact, any population prospect axiology that satisfies transitivity and State-Wise Dominance—faces additional problems due to the introduction of uncertainty, even when the scope is limited to non-trivial prospects, as we shall soon see.

But first, to appreciate some of the problems that non-total Utilitarian theories are faced with in the risk-free setting, let's look more closely at Average Utilitarianism, according to which the value of a population is found by summing up the welfare of the people in the population and dividing this sum with the number of people in the population.²² (After examining Average Utilitarianism and its prospect extension, we shall briefly discuss Critical Level Utilitarianism and its prospect extension.) Hence, the contributive value of a person's welfare depends on the average welfare of the population of which it is a member: The contributive value of a person's welfare is positive (negative, neutral) if the person has a higher (lower, same) welfare than the average welfare in the population.

Average Utilitarianism (AU) For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$:

$$\mathbf{x} \lesssim \mathbf{y} \Leftrightarrow AU(\mathbf{x}) = \bar{x} \leq AU(\mathbf{y}) = \bar{y}$$

Now, we can extend this theory to situations of risk—thus creating Prospect Average Utilitarianism—in an analogous way to how Total Utilitarianism was extended to Prospect Total Utilitarianism:

Prospect Average Utilitarianism (PAU) For any $P, Q \in \mathbf{P}$:

$$P \lesssim Q \Leftrightarrow \sum_i p(s_i)AU(P(s_i)) \leq \sum_i p(s_i)AU(Q(s_i))$$

Average Utilitarianism gives rise to the *Sadistic Conclusion*, which informally says that it is better to add to a population people with negative rather than positive welfare.²³ More formally:

²⁰ See, e.g., Parfit (1984); Arrhenius (2000b, forthcoming).

²¹ The same holds for the recent prospect axiology presented in Asheim and Zuber (2016) which is an extension of the population axiology in Asheim and Zuber (2014). The latter violates some compelling adequacy conditions, such as the Non-Extreme Priority Condition (Arrhenius forthcoming).

²² The problem with Average Utilitarianism that we focus on here also arises for Critical Level Utilitarianism. See Arrhenius (2000a, 2000b, forthcoming).

²³ See, e.g., Arrhenius (2000b).

Sadistic Conclusion There is an $\mathbf{x} \in \mathbf{X}$, a $\mathbf{y} \in \mathbf{X}^+$, and a $\mathbf{z} \in \mathbf{X}^-$ such that $(\mathbf{x}, \mathbf{y}) \prec (\mathbf{x}, \mathbf{z})$.

Even worse, Average Utilitarianism violates the seemingly unassailable *Weak Non-Sadism*, which is even logically weaker than what is required to avoid the Sadistic Conclusion. Informally, Weak Non-Sadism implies that there is a negative welfare level and a number of people at this level such that an addition of any number of people with positive welfare is at least as good as an addition of the negative welfare people. More formally:

Weak Non-Sadism There is a $z \in \mathbb{R}^-$ and an $n \in \mathbb{N}$, such that for any $\mathbf{x} \in \mathbf{X}$, for any $y \in \mathbb{R}^+$, and for any $m \in \mathbb{N}$, $(\mathbf{x}, (z)_n) \lesssim (\mathbf{x}, (y)_m)$.

Prospect Average Utilitarianism of course violates Weak Non-Sadism too, that is, for trivial prospects. Moreover, it is straightforward to verify that Prospect Average Utilitarianism also violates, *for non-trivial prospects*, a probabilistic analog of Weak Non-Sadism that is implied by Weak Non-Sadism together with the aforementioned dominance condition:

Probabilistic Weak Non-Sadism There is a $z \in \mathbb{R}^-$ and an $n \in \mathbb{N}$, such that for any $\mathbf{x}, \mathbf{x}' \in \mathbf{X}$, any $y \in \mathbb{R}^+$, any $m \in \mathbb{N}$, any $P, Q \in \mathbf{P}$, and any $E \subseteq \mathbf{E}$, if $P(E) = (\mathbf{x}, (z)_n)$ and $P(\neg E) = (\mathbf{x}, \mathbf{x}')$, while $Q(E) = (\mathbf{x}, (y)_m)$ and $Q(\neg E) = (\mathbf{x}, \mathbf{x}')$, then $P \lesssim Q$.

Perhaps more interestingly, while Average Utilitarianism (and Critical Level Utilitarianism) of course avoids the risk-free Repugnant Conclusion, Prospect Average Utilitarianism (and Prospect Critical Level Utilitarianism as well as of course Prospect Total Utilitarianism) leads to what we might call the *Risky Repugnant Conclusion*²⁴:

Risky Repugnant Conclusion For any $E \subset \mathbf{E}$, for any $\mathbf{y}, \mathbf{z} \in \mathbf{X}$, there is an $\mathbf{x} \in \mathbf{X}$, such that if $P(\neg E) = \mathbf{x}$ and $P(E) = \mathbf{z}$ while trivial prospect Q results in \mathbf{y} for sure, then $Q \prec P$.

The above observation means that while Average Utilitarianism (and Critical Level Utilitarianism) avoids the Repugnant Conclusion, *Prospect Average Utilitarianism* (and Prospect Critical Level Utilitarianism) recommends prospects that will *almost certainly* result in a Repugnant Conclusion, that is, in the special case where \mathbf{z} is a huge population consisting of people who have very low positive welfare while \mathbf{y} is a smaller population where everyone is very well-off. So, again we see that a prospect axiology leads to the probabilistic analogue of the conclusion that the corresponding risk-free axiology is celebrated for avoiding.

²⁴ McCarthy et al. (2020) discuss what one could think of as another probabilistic variant of the Repugnant Conclusion. However, the variant that they discuss can be avoided by Critical Level Utilitarianism, unlike the risky variant that we discuss.

For the sake of completeness, and to convince the reader that Prospect Critical Level Utilitarianism really does lead to the Risky Repugnant Conclusion, it is worth formally stating the theory. The general intuition behind Critical Level Utilitarianism is that only people with welfare that is above some “critical level”, which is typically assumed to be higher than the neutral level, can increase the value of a population. More precisely, the risk-free version of the theory states that for some critical level $c \geq 0$:²⁵

Critical-Level Utilitarianism (CLU) For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$:

$$\mathbf{x} \lesssim \mathbf{y} \Leftrightarrow CLU(\mathbf{x}) = \sum_i [x_i - c] \leq CLU(\mathbf{y}) = \sum_i [y_i - c]$$

Blackorby et al. (2005) have extended Critical Level Utilitarianism²⁶ to a prospect theory by essentially applying Expected Utility Theory to population prospects where each population is evaluated by the above formula. That is, they extend Critical Level Utilitarianism in the same way as we have extended Average and Total Utilitarianism (albeit using a somewhat different notation and formal framework than we do). The resulting theory states that:

Prospect Critical-Level Utilitarianism For any $P, Q \in \mathbf{P}$:

$$P \lesssim Q \Leftrightarrow \sum_i p(s_i) CLU(P(s_i)) \leq \sum_i p(s_i) CLU(Q(s_i))$$

It should be evident that while Critical Level Utilitarianism was partly designed to avoid the Repugnant Conclusion, Prospect Critical Level Utilitarianism leads to the Risky Repugnant Conclusion. For as long as the people in the \mathbf{x} population are sufficiently high above the critical level, then Prospect Critical Level Utilitarianism will instruct us to gamble on \mathbf{x} rather than getting \mathbf{y} for sure, irrespective of how much more likely the gamble is to result in \mathbf{z} than in \mathbf{x} and irrespective of how much worse off people are in \mathbf{z} compared to in \mathbf{y} .

Now, average and critical level utilitarians could avoid the Risky Repugnant Conclusion by postulating that there is a limit to how high a person’s welfare can be. But what we take this and the previous result to illustrate—and the point we want to emphasise—is that new problems appear when we try to turn a population axiology into a population prospect axiology. For, as we have seen, natural formulations of each of the three standard versions of Utilitarianism lead to prospect-versions of the conclusions that they avoid in the risk-free settings, as long as they do not put an upper bound on the moral value of a population, for instance, by setting an upper

²⁵ It is worth noting that c need not be a constant, but can be a function of the size of the evaluated population. A generalisation of Critical Level Utilitarianism was introduced by Blackorby and Donaldson (1984) and subsequently explored in depth by Blackorby et al. (2005) and Bossert (2017). The generalisation, which we ignore, adds a function on $x_i - c$ that is non-decreasing and non-convex (so, weakly or strictly concave or linear).

²⁶ That is, the generalised version mentioned in footnote 25.

limit on possible individual welfare and population size (in the case of Average Utilitarianism, the former limit suffices).²⁷

Here is a more general way of putting the last observation. The population axiologies we have considered all tell us that the best population maximises some value; e.g., total welfare, average welfare, total welfare above critical level. The focus on this value helps the axiology in question avoid some problem in the risk-free setting; for instance, by focusing on maximising average welfare, Average Utilitarianism avoids the Repugnant Conclusion, while by focusing on total welfare, Total Utilitarianism avoids the Sadistic Conclusion. However, if for each of these theories, the value it wants us to focus on is unbounded from above, then any prospect axiology that is based on expected utility theory will recommend that we choose some prospect that has an arbitrarily high chance of resulting in a terrible population (as judged by that axiology), over a risk-free “prospect” that can only result in a wonderful population (as judged by that axiology). So, formulated as an (informal) impossibility result, the point is that there is no *expectational* population prospect axiology, whose value is unbounded from above, that avoids recommending a prospect that has an arbitrarily high chance of resulting in a terrible population even when a wonderful population is on offer.²⁸

We can in fact put the last observation as a (very simple) formal and general theorem.²⁹ Recall that the above results hold for axiologies such as Total Utilitarianism and Average Utilitarianism according to which moral value is unbounded (from above). So, now let $V : \mathbf{X} \rightarrow \mathbb{R}$ be a generic (i.e., could be totalist, could be average, etc.) moral value function that is unbounded from above, that is, for any $k \in \mathbb{R}$, there is an $\mathbf{x} \in \mathbf{X}$ such that $V(\mathbf{x}) > k$.

Expected Value Theorem For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$ and for any $p \in (0, 1)$, there is a $\mathbf{z} \in \mathbf{X}$ such that

$$pV(\mathbf{y}) + (1 - p)V(\mathbf{z}) > V(\mathbf{x})$$

Proof Since V is unbounded from above, we can always find some \mathbf{z} such that

$$V(\mathbf{z}) > \frac{V(\mathbf{x}) - pV(\mathbf{y})}{(1 - p)}$$

□

To see why the result that Prospect Average Utilitarianism implies the Risky Repugnant Conclusion is an instance of the Expected Value Theorem, note that the

²⁷ However, note that even in the non-risky case, Average Utilitarianism implies repugnant-like conclusions. See e.g., Arrhenius (2000a, forthcoming), Budolfson and Spears (2018).

²⁸ We thank a referee for inspiring this way of putting the issue and for making us see the need of including the formal and general theorem, instead of only including the above instances.

²⁹ The result that follows could of course be generalised even further, such that it doesn't apply only to populations and population prospects but to any outcomes and any prospects. This should be evident since the result in no way depends on \mathbf{X} being a set of populations.

average utilitarian moral value function is an instance of V , and let \mathbf{x} be any population where all members have very high welfare, let \mathbf{y} be some population where all members have very low but still positive welfare, and choose some \mathbf{z} with a sufficiently high average welfare. Similar reasoning shows that the result that Prospect Total Utilitarianism implies the Risky Very Sadistic Conclusion is an instance of the Expected Value Theorem, and so on.

Now, the Expected Value Theorem could be seen as a motivation for seeking population axiologies that put an upper bound on moral value.³⁰ We are personally sceptical of such theories, for reasons that we do not have the space to discuss here. Still, we acknowledge that the results of this paper might provide some reason to rethink that scepticism.³¹

For the remainder of this paper, the question we will focus on is whether the impossibility results that have been proven for risk-free population axiology—that is, results showing that no population axiology can satisfy different sets of adequacy conditions (i.e., axioms) that many find to be intuitively compelling—also hold for non-trivial prospects. We know from the study of risk-free population axiology that there is a great advantage to focusing on exploring the logical consistency of different adequacy conditions, rather than formulating new theories in response to the impossibility results that have been found for the old theories. After all, that search has, so far, only resulted in new impossibility results (as reviewed in Arrhenius, [forthcoming](#)).

In the spirit of the axiomatic approach, we shall, in the next section, formulate a weak and very plausible state-dominance condition that some would see as ensuring that the ranking of population prospects coheres with the ranking of populations. As we show in section 4, this condition suffices to generate impossibility results for population prospect axiology that are similar to the traditional impossibility results for (risk-free) population axiology. Moreover, as we shall explain in section 4, some condition such as that introduced in the next section is *needed* to generate impossibility results for *non-trivial* population prospects, since the traditional adequacy conditions for risk-free population axiology say nothing whatsoever about how to rank non-trivial prospects.

Nevertheless, one might have already suspected that since any prospect axiology should satisfy some condition of continuity in probabilities, the impossibility results for population axiology would reappear in population prospect axiology. For instance, one might think that if a risk-free adequacy condition for population axiology implies that \mathbf{x} is better than \mathbf{y} , then such a continuity condition for population prospect axiology would ensure that a non-trivial prospect that results in \mathbf{x} with a probability that is arbitrarily close to 1 is better than a prospect that results in \mathbf{y} with a probability that is arbitrarily close to 1, irrespective of what other populations the two prospects could result in.

³⁰ This was pointed out by a referee.

³¹ But note that an upper bound on the population size would not be sufficient, as we also would need an upper bound on individual welfare to get bounded moral value.

The above application of continuity is one simple way of generating population prospects impossibility results on the back of the traditional impossibility results for risk-free prospects. For the remainder of this paper, we will however explore different approaches, that do not assume continuity in probabilities. Instead, we assume (except for in the final result) a dominance condition which some will find less questionable than continuity, and which the next section formally introduces.

4 Population Prospect Impossibility Theorems

As an illustration of a straightforward comparison of population prospects, consider prospects P_1 and P_2 in Table 1. We assume, in the below example (and indeed throughout this paper), that the only axiologically relevant features at stake are, first, the welfare of the people in the populations that the prospects might result in, and, second, the probabilities that the prospects result in these populations. Moreover, in the example below, we assume that the only person whose welfare P_1 and P_2 affect is Ann, who will lead a horrible and suffering life (denoted by “Ann: -100”) if we choose P_2 . The same is true if we choose P_1 and state of the world s_1 obtains. However, in state s_2 the choice of P_1 means that Ann will lead an excellent life (which we denote by “Ann: 100”).

It seems evident that P_1 is better than P_2 . After all, Ann is sure to lead a suffering life given the choice of P_2 , but she could be spared these sufferings and lead an excellent life given the choice of P_1 .

The above reasoning is an instance of the well-known *State-Wise Dominance* principle from decision theory. The role of this principle could be seen as ensuring that the ranking of prospects coheres with the ranking of the outcomes (in our case, the populations) that the prospects could result in.³² We see no reason why one would reject this as an adequacy condition for population prospect axiologies, assuming, as we have done, that the only axiologically relevant feature at stake is the welfare of the people involved and the probabilities of the possible populations (Fleurbaey 2010, 654 makes a similar point).³³ Nevertheless, one of the results we present in the next section does not assume State-Wise Dominance in full generality; and one result does not assume the principle at all.³⁴

Let us call the fully general implication that State-Wise Dominance has for population prospects *Population State Dominance*, and formulate it thus:

³² The classical theories of rational choice under uncertainty—such as Ramsey’s 1926/1990, Savage’s 1954, and Jeffrey’s 1965—all impose State-Wise Dominance for the type of situations we are considering, i.e., situations where the states’ probabilities are unaffected by the agents’ choice.

³³ So, for instance, issues of fair chances don’t arise (see, e.g., Stefánsson 2015.)

³⁴ It might be worth mentioning that unlike *Event-Wise Dominance*, State-Wise Dominance does not rule out intuitively reasonable risk attitudes that are inconsistent with standard Expected Utility Theory, such as the so-called *Allais Preference* (named after Allais 1953).

Table 1 Illustration of dominance relationship

	s_1	s_2
P_1	Ann: -100	Ann: 100
P_2	Ann: -100	Ann: -100

Population State Dominance If for all $s_i \in \mathbf{S}$, $P_2(s_i) \lesssim P_1(s_i)$, then $P_2 \lesssim P_1$; and if, in addition, for at least one non-null³⁵ s_j , $P_2(s_j) < P_1(s_j)$, then $P_2 < P_1$.

We now turn to the question: is it possible to formulate a population prospect axiology that is unburdened by (something like) the traditional impossibility results for population axiology? Although the traditional (risk-free) adequacy conditions are known to be mutually inconsistent as conditions on how to rank populations, they are not mutually inconsistent as conditions on how to rank *non-trivial* prospects, since none of them says anything about how to rank such prospects. Hence, a condition like Population State Dominance is needed to connect the ranking of outcomes to the ranking of non-trivial prospects.

We shall focus on three adequacy conditions for (risk-free) population axiology that generate a very simple and easy to demonstrate impossibility theorem, which we will extend to population prospect axiology. Some of these conditions can be and have been questioned. For instance, one might want to replace Quality by weaker a weaker condition that does not directly rule out the Repugnant Conclusion. However, as will become evident, the various other impossibility theorems that have been proven with logically weaker and intuitively more compelling adequacy conditions for population axiology, for instance by weakening Quality (see, e.g., Arrhenius 2011, [forthcoming](#)), can be extended to population prospect axiology in a way analogous to how we extend the simple theorem. So, the aim here is not to prove an impossibility theorem involving only intuitively compelling adequacy conditions. The aim is instead to describe a recipe for translating *any* theorem for (risk-free) population axiology into a theorem for prospect population axiology on a domain that contains *only* non-trivial prospects. (Later we prove a general theorem that establishes that this recipe can indeed be used to extend any axiology-result to a prospect-result.)

Below we make use of one more piece of notation. We stipulate that there is a $\gamma \in \mathbb{R}^+$ —which can be arbitrarily close to 0—such that for any $x, y \in \mathbb{R}^+$ if $y < x$ but $x - y \leq \gamma$, then y is “slightly lower” than x . The three adequacy conditions for risk-free population axiology that we focus on are:

Egalitarian Dominance For any $\mathbf{x}, \mathbf{y} \in \mathbf{X}$, if $n(\mathbf{x}) = n(\mathbf{y})$ and for all $i, j \in \mathbb{N}$, $y_i < x_j$ and $x_i = x_j$, then $\mathbf{y} < \mathbf{x}$.

Quantity For any $x, y \in \mathbb{R}^+$ such that y is slightly lower than x , and for any $n \in \mathbb{N}$, there is an $m \in \mathbb{N}$, such that $m > n$, and such that $(x)_n \lesssim (y)_m$.

³⁵ A state is non-null if it has some probability of obtaining (in this case, according to the social planner).

Quality There is an $x \in \mathbb{H}$ and an $n \in \mathbb{N}$, such that for any $y \in \mathbb{L}$ and any $m \in \mathbb{N}$, $(y)_m \succsim (x)_n$.

Egalitarian Dominance is as uncontroversial as a principle of population ethics gets. Quality is a condition that rules out theories that imply the Repugnant Conclusion. Quantity should appeal to those who find some truth “the more of the good, the better”.

It turns out that the above seemingly weak conditions are not mutually compatible, as illustrated by the following simple and well-known impossibility theorem (for risk-free population axiology):

Simple Impossibility Theorem (SIT) No social welfare relation \succsim on \mathbf{X} satisfies Egalitarian Dominance, Quality, and Quantity.³⁶

To see why the above theorem is true, consider the sequence in Fig. 1. We assume that well-being is perfectly equally distributed in each population. Moreover, assume that \mathbf{x}_1 in the diagram above is a population with very high welfare and that \mathbf{y} is a population with very low positive welfare (the width of a block represents the number of people in the population, the height represents their welfare; dashes indicate that the block in question should really be much wider than shown). According to Quantity, there is a population \mathbf{x}_2 with slightly lower welfare than \mathbf{x}_1 but which is at least as good as \mathbf{x}_1 ; a population \mathbf{x}_3 with slightly lower welfare than \mathbf{x}_2 but which is at least as good as \mathbf{x}_2 ; and so forth. Hence, we will finally reach population \mathbf{x}_r with very low positive welfare.³⁷ Now, recall that we assume the social welfare relation to be transitive. Hence, \mathbf{x}_r is at least as good as \mathbf{x}_1 . But \mathbf{y} is better than \mathbf{x}_r according to Egalitarian Dominance. Hence, by transitivity, \mathbf{y} is better than \mathbf{x}_1 . However, by Quality, \mathbf{x}_1 is at least as good as \mathbf{y} (for some \mathbf{x}_1 and \mathbf{y} that can be used in an argument with the above structure). So, we get a contradiction: \mathbf{y} is better than \mathbf{x}_1 which is at least as good as \mathbf{y} . Hence, the Simple Impossibility Theorem: there is no social welfare relation—in other words, no population axiology—which satisfies Egalitarian Dominance, Quality, and Quantity.

The Simple Impossibility Theorem can easily be extended to the following impossibility result for population prospect axiologies, even if we limit the domain of the social welfare relation to prospects that are non-trivial (in the sense described above):

³⁶ This theorem was first presented in Arrhenius (2000a). See also Arrhenius (2016)

³⁷ Here we exploit the fact that the formal framework we have been employing entails what has been called *Finite Fine-grainedness*: There exists a finite sequence of slight welfare differences between any two welfare levels. Finite Fine-grainedness ensures that one can get from one welfare level to another in a finite number of steps of intuitively slight welfare difference. Examples of such welfare differences could be some minor pain or pleasure or a shortening of life by a minute or two. For further discussion of Finite Fine-grainedness, and of possible theories of welfare that violate this assumption, see Arrhenius (2000a, 2005, 2016, forthcoming), Arrhenius and Rabinowicz (2015). For interesting efforts to challenge Finite Fine-grainedness (in light of the impossibility theorems in population ethics), see Thomas (2018) and Carlson (2022). For a very recent impossibility result for prospects without Finite Fine-grainedness, see Thornley (2021).

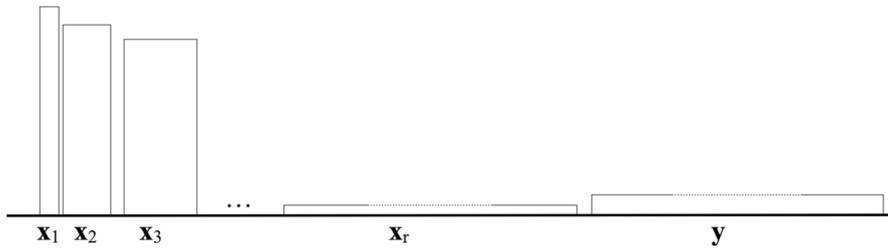


Fig. 1 Illustration of SIT

Simple Prospect Impossibility Theorem (SPIT) No social welfare relation \preceq on the set of non-trivial prospects, \mathbf{P}_{nt} ,³⁸ satisfies Egalitarian Dominance, Quality, Quantity, and Population State Dominance.

The prospects in Table 2 establish SPIT. P_1 to P_{r+1} are different population prospects, whereas s_1 and s_2 are different states of the world that determine which population results from the choice of each prospect. We assume that P_1 to P_{r+1} are non-trivial, meaning that the probabilities of states s_1 and s_2 are strictly between zero and one. Now, assume that $\mathbf{x}_1, \mathbf{x}_2, \dots$, and \mathbf{y} in Table 2 are the populations depicted in Fig. 1. So, for each \mathbf{x}_i , \mathbf{x}_{i+1} is at least as good as \mathbf{x}_i according to Quantity, \mathbf{y} is better than \mathbf{x}_r , according to Egalitarian Dominance, and thus \mathbf{y} is better than \mathbf{x}_1 by transitivity. However, \mathbf{x}_1 is at least as good as \mathbf{y} according to Quality. But that means that according to Population State Dominance, P_{r+1} is better than P_1 and P_1 is at least as good as P_{r+1} , irrespective of what population \mathbf{z} represents. So, again we reach a contradiction. Hence, the Simple Prospect Impossibility Theorem: there is no social welfare relation—and thus no population *prospect* axiology—on a domain of non-trivial prospects which satisfies Egalitarian Dominance, Quality, Quantity, and Population State Dominance. (It is straightforward to see how the result can be extended to multiple non-null states.)

Note that since both Quality and Quantity are stated with the weak preference relation—that is, they are claims about the “at least as good as” relation, to use axiological terms—we need a version of Population State Dominance that applies dominance reasoning even when there is only “weak” state-wise dominance (i.e., where there is only weak preference for one prospect over another given some state). However, by strengthening Quality, we can make do with a weaker dominance condition that only applies dominance reasoning when there is strict state-wise dominance.³⁹

We could of course also replace Population State Dominance with *Stochastic Dominance*, which, in a state-space framework like ours, entails Population State Dominance. Finally, it might also be worth pointing out that the assumption that the different prospects result in the *same* population if one of the two states obtains is

³⁸ Of course, there is no such relation on the set of trivial prospects ($\mathbf{P} \setminus \mathbf{P}_{nt}$) either; but that we already knew from SIT.

³⁹ We thank Dean Spears for encouraging us to clarify this.

Table 2 Prospects that prove SPIT

	s_1	s_2
P_1	\mathbf{z}	\mathbf{x}_1
P_2	\mathbf{z}	\mathbf{x}_2
\cdot	\cdot	\cdot
P_r	\mathbf{z}	\mathbf{x}_r
P_{r+1}	\mathbf{z}	\mathbf{y}

not essential to the above result (nor the next theorems we introduce). In fact, we can prove the following general theorem, which implies SPIT as a special case.⁴⁰

Dominance Theorem Let \mathcal{A} be a set of adequacy conditions that no social welfare relation on the set of trivial prospects, \lesssim on $\mathbf{P} \setminus \mathbf{P}_{nt}$, satisfies. Then no social welfare relation \lesssim on \mathbf{P}_{nt} satisfies \mathcal{A} and Population State Dominance.

Proof For any non-null state $s_i \in \mathbf{S}$, let \mathcal{P} denote the set of all prospects that map $\mathbf{S} \setminus \{s_i\}$ to some particular population vector $\mathcal{V} = (\mathbf{x}, \dots, \mathbf{y})$. Now let \mathcal{P}_x be the prospect in \mathcal{P} that results in population \mathbf{x} in state s_i and similarly for prospect \mathcal{P}_y and population \mathbf{y} . Now let P_x be the trivial prospect that results in population \mathbf{x} for sure, and similarly for P_y and population \mathbf{y} . From Population State Dominance it follows that $\mathcal{P}_x \lesssim \mathcal{P}_y$ just in case $P_x \lesssim P_y$. Now, note that $P_x, P_y \in \mathbf{P} \setminus \mathbf{P}_{nt}$. Hence, since by assumption, the relation \lesssim on $\mathbf{P} \setminus \mathbf{P}_{nt}$ fails to satisfy the conditions in \mathcal{A} , some pair of trivial prospects P_x, P_y establishes that \lesssim on \mathbf{P}_{nt} cannot satisfy both \mathcal{A} and Population State Dominance. □

Before introducing the next theorem, it might be worth reminding the reader of a point that we previously made, namely, that prospect impossibility results such as those proven above cannot be achieved by only using risk-free adequacy conditions, such as Egalitarian Dominance, Quantity, and Quality. For while these risk-free conditions tell us how to rank *populations* (and hence how to rank *trivial* population prospects), they tell us nothing about how to rank *non-trivial* prospects that could result in different populations. For instance, recall that Egalitarian Dominance, Quality, and Quantity together entail that \mathbf{y} is better than \mathbf{x}_1 and that \mathbf{x}_1 is at least as good as \mathbf{y} , that is, the conditions entail an inconsistent ranking of *populations* (or risk-free outcomes). But, in and of itself, that does not suffice to get an impossibility result for non-trivial prospects, since this non-transitive outcome-ranking by itself does not preclude that there exists a consistent (in particular, transitive) ranking of *non-trivial prospect*. To take a simple (but of course implausible) example, suppose that the only adequacy condition on non-trivial prospects says that all non-trivial prospects are equally preferable. That would not, of course, remove the impossibility

⁴⁰ We thank a referee for encouraging us to include the general theorem rather than only the special cases, and for suggesting something close to this argument.

result for populations (i.e., risk-free outcomes). But it would mean that the “set” of adequacy conditions on non-trivial prospects is consistent.

So, if a person were to maintain that a decision-maker is never faced with choices between populations, only with choices between population *prospects*, then they could argue that from a practical (decision-making) point of view, the inconsistency generated by Egalitarian Dominance, Quantity, and Quality is not so worrying. However, if this person accepts Population-State Dominance for prospects, and moreover accept Egalitarian Dominance, Quantity, and Quality for populations, then they would have to admit that it is worrying (even from a practical point of view) that Population-State Dominance results in an inconsistent ranking of prospects when the ranking of the possible outcomes (i.e., populations) satisfies Egalitarian Dominance, Quantity, and Quality.

Another way to derive an impossibility result for non-trivial prospects, without using a condition like Population State Dominance to connect the ranking of outcomes to the ranking of prospects, is to reformulate the standard conditions as prospect conditions, a strategy we next explore. Consider first the following, where state-dominance is essentially built into each condition:

Egalitarian Prospect Dominance For any $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbf{X}$ such that $n(\mathbf{x}) = n(\mathbf{y})$ and $y_i < x_j = x_i$ for all $i, j \in \mathbb{N}$, for any $E \subseteq \mathbf{E}$, and for any $P, Q \in \mathbf{P}$, if $P(\neg E) = \mathbf{x}$ and $P(E) = \mathbf{z}$ while $Q(\neg E) = \mathbf{y}$ and $Q(E) = \mathbf{z}$, then $Q \prec P$.

Quantity Prospect Dominance For any $\mathbf{z} \in \mathbf{X}$ and for any $x, y \in \mathbb{R}^+$ such that y is slightly lower than x , and for any $n \in \mathbb{N}$, there is an $m \in \mathbb{N}$, such that $m > n$, and such that for any $E \subseteq \mathbf{E}$ and for any $P, Q \in \mathbf{P}$, if $P(E) = (x)_n$ and $P(\neg E) = \mathbf{z}$, while $Q(E) = (y)_m$ and $Q(\neg E) = \mathbf{z}$, then $P \lesssim Q$.

Quality Prospect Dominance There is an $x \in \mathbb{H}$ and an $n \in \mathbb{N}$, such that for any $y \in \mathbb{L}$, any $m \in \mathbb{N}$, any $\mathbf{z} \in \mathbf{X}$, any $E \subseteq \mathbf{E}$, and any $P, Q \in \mathbf{P}$, if $P(E) = (x)_n$ and $P(\neg E) = \mathbf{z}$, while $Q(E) = (y)_m$ and $Q(\neg E) = \mathbf{z}$, then $Q \lesssim P$.

Now, it should be straightforward to see that even if the domain is restricted to non-trivial prospects, that is, prospects that confer a probability strictly between zero and one on different populations, the above three conditions form an inconsistent set. That is, we can prove:

Pure Prospect Impossibility Theorem (PIT) No social welfare relation \lesssim on \mathbf{P}_{nt} satisfies Egalitarian Prospect Dominance, Quality Prospect Dominance, and Quantity Prospect Dominance.

The prospects in Table 2, for instance, establish the Pure Prospect Impossibility Theorem. For notice that by transitivity, Egalitarian Prospect Dominance, Quality Prospect Dominance, and Quantity Prospect Dominance result in an inconsistent ranking of the prospects in Table 2, in exactly the same way as the original

Egalitarian Dominance, Quality, Quantity and Population State Dominance resulted in an inconsistent ordering of the prospects in Table 2.

One advantage of PIT, as compared to SPIT, is that PIT does not assume Population State Dominance to always hold. So, in particular, those who want to resist Population State Dominance due concern for fairness in the distribution of chances (e.g., Diamond 1967, Stefánsson 2015), might be more troubled by the impossibility identified by PIT, which only assumes that Population State Dominance holds with respect to the populations involved in the conditions.

However, in case someone objects to *any* appeal to Population State Dominance, even a restricted one as used in the Pure Prospect Impossibility Theorem, it might be worth noting that impossibility results can be generated without any such appeal, albeit with more complex conditions.⁴¹

Consider first Probability-Quantity Tradeoff, the thought behind which is that we can reduce the probability that a prospect results in the population $(x)_m$, where people have very high welfare, without making the resulting prospect worse, as long as we sufficiently increase the number of people in the alternative population that the prospect results in if it does not result in $(x)_m$:

Probability-Quantity Tradeoff For any $m, n \in \mathbb{N}$, any $x \in \mathbb{H}$ and $y \in \mathbb{L}$, for any $E \in \mathbf{E}$ such that $p(E) = q > 0$ and for any $P \in \mathbf{P}$ such that $P(E) = (x)_m$ and $P(\neg E) = (y)_n$, there is some $F \in \mathbf{E}$, which possibly partly overlaps with E , such that $p(F) = q' < q$, and some $o \in \mathbb{N}$, such that if $Q(F) = (x)_m$ and $Q(\neg F) = ((y)_n, (y)_o)$, then $P \lesssim Q$.

Less formally: For any high welfare population and any low welfare population, and any prospect that with probability q results in the high welfare population and with probability $1 - q$ results in the low welfare population, there is another at least as good prospect that with probability $q' < q$ results in the high welfare population and with probability $1 - q'$ results in the union of the low welfare population with some additional lives at that same low level of welfare.

Next consider Probability-Quality Tradeoff, the thought behind which is that there is a limit to how much average welfare we are willing to risk foregoing for the sake of a chance at higher total welfare:

Probability-Quality Tradeoff There is an $x \in \mathbb{H}$ and an $n \in \mathbb{N}$, such that for any $k, l \in \mathbb{N}$, and for any $y \in \mathbb{L}$, there are some (possibly overlapping) events, $G, H \in \mathbf{E}$, such that $p(H) < p(G)$, such that if $P(G) = (x)_n$ and $P(\neg G) = (y)_k$, while $Q(H) = (x)_n$ and $Q(\neg H) = ((y)_k, (y)_l)$, then $Q \prec P$.

In plain English: There is some high welfare population, such that for any two disjoint low welfare populations, there are probabilities p and q , where $p > q$, such that a prospect that with probability p results in the high welfare population and with probability $1 - p$ results in one of the low welfare populations is better

⁴¹ We thank Tim Campbell for suggesting we add a result like this to the paper.

than any prospect that with probability q results in the high welfare population and with probability $1 - q$ results in the union of the two low welfare populations.

As can easily be verified, Prospect Average Utilitarianism violates Probability-Quantity Tradeoff whereas Prospect Total Utilitarianism violates Probability-Quality Tradeoff.

To see how Probability-Quantity Tradeoff and Probability-Quality Tradeoff jointly lead to inconsistency, given transitivity, suppose we start with a probability p that is close to 1. (To make the argument as simple as possible, we replace the events/states by their probabilities.) By Probability-Quantity Tradeoff, there is some quantity δ that we could subtract from p such that the following holds. Suppose that prospect P_0 has probability p of resulting in a population $(x)_m \in \mathbb{H}^m$ and probability $1 - p$ of resulting in some population $(y)_n \in \mathbb{L}^n$. Moreover, suppose that prospect P_1 has probability $p - \delta$ of resulting in population $(x)_m$ and probability $1 - (p - \delta)$ of resulting in $((y)_n, (y)_o)$. Then P_1 is no worse than P_0 . But continuing this reasoning, we can generate a sequence of prospects, by incrementally increasing the quantity that is subtracted from p , until we reach a quantity γ for which the following holds. Suppose that P_n has probability $p - \gamma$ of resulting in $(x)_m$ and probability $1 - (p - \gamma)$ of resulting in some population $((y)_n, (y)_o, \dots, (y)_p)$, which is much bigger than $(y)_n$. Then P_n is no worse than the immediately preceding prospect in the sequence, and the same is true of any other prospect in the sequence. Hence, by transitivity, P_n is no worse than P_0 . Since we get this result irrespective of the welfare (and size) of the $(x)_m$ population, and irrespective of p and γ , we have a violation of Probability-Quality Tradeoff.

5 Concluding Remarks

We have found that the search for a population axiology free from constraintive implications is not made easier by the introduction of risk. For instance, we have shown how any impossibility theorem that has been proven for risk-free population axiology can be extended to population prospect axiology with the use of an instance of the traditional State-Wise Dominance axiom from decision theory. Moreover, we have seen that the introduction of risk into population ethics brings with it some new problems. Perhaps most interestingly, we have seen that natural extensions of traditional population axiologies lead to prospect-versions of the problems that they are praised for avoiding in the risk-free setting. In addition, we have seen that impossibility results can be derived from probabilistic axioms only (i.e., without any risk-free axioms). All of this illustrates that contrary to what one might have hoped, the task of finding an attractive axiology for non-trivial population prospects is even harder than the task of finding an attractive population axiology.

Funding Open access funding provided by Stockholm University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Allais M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'école Américaine. *Econometrica* 21(4):503–546
- Arrhenius G (2000a) Future generations: a challenge for moral theory. Uppsala University Printers
- Arrhenius G (2000) An impossibility theorem for welfarist axiologies. *Econ Philos* 16(2):247–266
- Arrhenius G (2003) Feldman's desert-adjusted utilitarianism and population ethics. *Utilitas* 15(2):225
- Arrhenius G (2003b) The very repugnant conclusion. *Uppsala Philosophical Studies* 51
- Arrhenius G (2005) Superiority in value. *Philos Stud* 123(1–2):97–114
- Arrhenius G (2009) Can the person affecting restriction solve the problems in population ethics? In: Roberts MA, Wasserman DT (eds) *Harming future persons*. Springer, Berlin, pp 289–314
- Arrhenius G (2011) The impossibility of a satisfactory population ethics. In: Dzhaferov E, Perry L (eds) *Descriptive and normative approaches to human behavior*. World Scientific, Singapore
- Arrhenius G (2016) Population ethics and different-number-based imprecision. *Theoria* 82(2):166–181
- Arrhenius G, Rabinowicz W (2015) Value superiority. In: Hirose I, Olson J (eds) *The Oxford handbook of value theory*. Oxford University Press, Oxford, pp 225–248
- Arrhenius G (forthcoming) *Population ethics: the challenge of future generations*. Oxford University Press, Oxford
- Asheim G, Zuber S (2014) Escaping the repugnant conclusion: rank-discounted utilitarianism with variable population. *Theor Econ* 9(3):629–650
- Asheim G, Zuber S (2016) Evaluating intergenerational risks. *J Math Econ* 65(C):104–117
- Bell D (1982) Regret in decision making under uncertainty. *Oper Res* 30:961–81
- Blackorby C, Donaldson D (1984) Social criteria for evaluating population change. *J Public Econ* 25(1–2):13–33
- Blackorby C, Bossert W, Donaldson D (1998) Uncertainty and critical-level population principles. *J Popul Econ* 11(1):1–20
- Blackorby C, Bossert W, Donaldson DJ (2005) *Population issues in social choice theory, welfare economics, and ethics*. Cambridge University Press, Cambridge
- Bossert W (2017) Anonymous welfarism, critical-level principles, and the repugnant and sadistic conclusions. Working paper, University of Montréal
- Broome J (1992) *Counting the cost of global warming*. White Horse Press, Stroud
- Broome J (2004) *Weighing lives*. Oxford University Press, Oxford
- Broome J (2010) The most important thing about climate change. In: Boston J, Bradstock A, Eng D (eds) *Public policy: why ethics matters*. ANU E Press, pp 101–116
- Broome J (2015) Climate change: life and death. In: Jeremy M (ed) *Climate change and justice*. Cambridge University Press, pp 184–200
- Buchak L (2013) *Risk and rationality*. Oxford University Press, Oxford
- Budolfson M, Spears D (2018) Why the repugnant conclusion is inescapable. Working paper, Princeton University CFI
- Carlson E (2022) On some impossibility theorems in population ethics. In: Krister B, Tim C, Elizabeth F-B (eds) *Gustaf Arrhenius. The Oxford handbook of population ethics*. Oxford University Press
- Diamond P (1967) Cardinal welfare, individualistic ethics, and interpersonal comparison of utility: comment. *J Polit Econ* 75
- Diecidue E, Somasundaram J (2017) Regret theory: a new foundation. *J Econ Theory* 172(C):88–119

- Feldman F (1997) *Utilitarianism, hedonism, and desert: essays in moral philosophy*. Cambridge University Press, Cambridge
- Fleurbaey M (2010) Assessing risky social situations. *J Polit Econ* 118(4):649–680
- Glover J (1977) *Causing death and saving lives*. Penguin Books, London
- Hammond PJ (1987) On reconciling Arrow's theory of social choice with Harsanyi's fundamental utilitarianism. In: Feiwel GR (ed) *Arrow and the foundations of the theory of economic policy*. Palgrave Macmillan, pp 179–221
- Hammond PJ (1988) Consequentialist demographic norms and parenting rights. *Social Choice Welf* 5(2/3):127–145
- Harsanyi JC (1953) Cardinal utility in welfare economics and in the theory of risk-taking. *J Polit Econ* 61(5):434–435
- Harsanyi JC (1955) Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *J Polit Econ* 63(4):309–321
- Hurka T (1983) Value and population size. *Ethics* 93(3):496–507
- Jeffrey R (1965) *The logic of decision*. The University of Chicago Press, Chicago
- Kahneman D, Tversky A (1979) Prospect theory: an analysis of decision under risk. *Econometrica* 47(2):263–291. <https://doi.org/10.2307/1914185>
- Loomes G, Sugden R (1982) Regret theory: an alternative theory of rational choice under risk. *Econ J* 92:805–824
- McCarthy D, Mikkola K, Thomas T (2020) Utilitarianism with and without expected utility. *J Math Econ* 87:77–113
- Nebel JM (2017) Priority, not equality, for possible people. *Ethics* 127(4):896–911
- Nebel JM (2019) An intrapersonal addition paradox. *Ethics* 129(2):309–343
- Ng Y-K (1989) What should we do about future generations?: Impossibility of Parfit's theory X. *Econ Philos* 5(2):235–253
- Parfit D (1984) *Reasons and persons*. Oxford
- Parfit D (2016) Can we avoid the repugnant conclusion? *Theoria* 82(2):110–127
- Pivato M (2020) Rank-additive population ethics. *Econ Theory* 69(4):861–918
- Quiggin J (1982) A theory of anticipated utility. *J Econ Behav Organ* 3(5):323–343
- Ramsey FP (1926/1990) *Truth and probability*. In: Mellor DH (ed) *Philosophical papers*. Cambridge University Press
- Roberts MA (2007) The non-identity fallacy: harm, probability and another look at Parfit's depletion example. *Utilitas* 19(3):267–311
- Roberts M (2010) The nonidentity problem. In: *Stanford encyclopedia of philosophy*
- Savage LJ (1954) *The foundations of statistics*. Wiley Publications in Statistics, New York
- Sider TR (1991) Might theory X be a theory of diminishing marginal value? *Analysis* 51(4):265–271
- Spears D, Budolfson M (2019) Why variable-population social orderings cannot escape the repugnant conclusion: proofs and implications. Discussion paper 12668, IZA
- Stefánsson HO (2015) Fair chance and modal consequentialism. *Econ Philos* 31(3):371–395
- Stefánsson HO, Bradley R (2015) How valuable are chances? *Philos Sci* 82(4):602–625
- Stefánsson HO, Bradley R (2019) What is risk aversion? *Br J Philos Sci* 70(1):77–102
- Thomas T (2018) Some possibilities in population axiology. *Mind* 127(507):807–832
- Thornley E (2021) The impossibility of a satisfactory population prospect axiology. *Philos Stud* 178(11):3671–3695
- Tversky A, Kahneman D (1992) Advances in prospect theory: cumulative representation of uncertainty. *J Risk Uncertain* 5(4):297–323
- von Neumann J, Morgenstern O (1944) *Theory of games and economic behavior*. Princeton University Press, Princeton
- Voorhoeve A, Fleurbaey M (2016) Priority or equality for possible people? *Ethics* 126(4):929–954
- Zuber S, Asheim GB (2012) Justifying social discounting: the rank-discounted utilitarian approach. *J Econ Theory* 147(4):1572–1601