Bounded Utilities and Ex Ante Pareto[*]

Petra Kosonen[†]

May, 2024

ABSTRACT: This paper shows that decision theories on which utilities are bounded, such as standard axiomatizations of Expected Utility Theory, violate Ex Ante Pareto if combined with an additive axiology, such as Total Utilitarianism. A series of impossibility theorems point toward Total Utilitarianism as the right account of axiology, while money-pump arguments put Expected Utility Theory in a favorable light. However, it is not clear how these two views can be reconciled. This question is particularly puzzling if utilities are bounded (as standard axiomatizations of Expected Utility Theory imply) because the total quantity of well-being might be infinite or arbitrarily large. Thus, there must be a non-linear transformation from the total quantity of well-being into utilities used in decision-making. However, such a transformation leads to violations of Ex Ante Pareto. So, the reconciliation of Expected Utility Theory and Total Utilitarianism prescribes prospects that are better for none and worse for some.

[*]For valuable feedback on this paper, I wish to thank Gustav Alexandrie, Jean Baccelli, Tomi Francis, Jake Nebel and the audiences of ANU Philosophy Departmental Seminar, the PWI conference 'Unbounded Utility in Ethics and Decision Theory', ISUS 2023, Axiology Workshop 2023 and the First Paris Conference: Frontiers of Economics and Philosophy. I especially want to thank Andreas Mogensen and Teruji Thomas for giving extensive feedback on this paper.

[†]Population Wellbeing Initiative, University of Texas at Austin. I would be grateful for comments: kosonenpetra@gmail.com

This paper investigates the compatibility of two standard theories: Total Utilitarianism and Expected Utility Theory with a bounded utility function. Let's call the combination of these views *Bounded Expected Totalism*. This paper shows that Bounded Expected Totalism violates *Ex Ante Pareto*, the principle that what is ex ante better for each individual is better overall.[1,2]

Ex Ante Pareto is often used by utilitarians to justify their theory in opposition to other views, such as prioritarianism and egalitarianism.[3] Insofar as Expected Utility Theory is the dominant theory of choice under uncertainty, the argument in this paper could be seen as undermining Total Utilitarianism. So, one might take the argument to support alternative population axiologies. On the other hand, the argument may speak differentially in favor of alternatives to bounded Expected Utility Theory, such as unbounded expected utility maximization.[4] Another alternative to Expected Utility Theory, discounting small probabilities, also leads to violations of Ex Ante Pareto.[5] The argument in this paper indirectly supports discounting small probabilities, as it shows that the plausibility of Ex Ante Pareto does not favor bounded Expected Utility Theory over discounting small probabilities.

The paper proceeds as follows. §1 introduces some background, such as why orthodox decision theory implies bounded utilities. §2 defines Bounded Expected

[1]There is no inconsistency with Harsanyi's social aggregation theorem. As will be explained later, a bounded expected totalist must reject Harsanyi's conclusion, so they cannot accept all his premises.

[2]This paper focuses on the compatibility of Expected Utility Theory and Total Utilitarianism, but the problem with Ex Ante Pareto arises for, for example, Critical-Level Utilitarianism in exactly the same way. The problem also arises for Average Utilitarianism and many other theories if individual utilities are unbounded. See for example the argument in Goodsell (2021), which applies to any axiology that is utilitarian in same-number cases. The contribution of this paper is showing that even if utilities are bounded, Total Utilitarianism combined with Expected Utility Theory violates Ex Ante Pareto.

[3]See Broome (1991, Ch 9), Parfit (1997), Rabinowicz (2002) and Fleurbaey (2018).

[4]See for example Wilkinson (2022).

[5]Kosonen (2021, §5). On discounting small probabilities and related topics, see Beckstead (2013, Ch 6), Smith (2014), Hájek (2014), Isaacs (2016), Monton (2019), Lundgren & Stefánsson (2020), Russell & Isaacs (2021), Wilkinson (2022), Kosonen (2022), Beckstead & Thomas (2023), Cibinel (2023) and Russell (2023).

Totalism more formally. §3 shows that Bounded Expected Totalism violates Ex Ante Pareto. §4 concludes by sketching how the examples relate to the classic result in this area, namely, Harsanyi's social aggregation theorem.

1 Background

This section begins by explaining Total Utilitarianism and Expected Utility Theory. Next, it shows how standard versions of Expected Utility Theory imply bounded utilities.

1.1 Total Utilitarianism and Expected Utility Theory

A series of impossibility theorems point toward Total Utilitarianism as the right account of axiology, while money-pump arguments put Expected Utility Theory in a favorable light as a theory of instrumental rationality.[6] According to Total Utilitarianism, a state of affairs is better than another just in case the total quantity of well-being it contains is greater. The total quantity of well-being in some state of affairs is the sum of individuals' well-being levels in that state of affairs.

> **Total Utilitarianism:** For all states of affairs A and B, A is at least as good as B if and only if the total quantity of well-being in A is at least as great as the total quantity of well-being in B.[7]

Expected Utility Theory, in turn, states that a prospect is better than another

[6] See Arrhenius (2000) and Gustafsson (2022*b*). The impossibility theorems point toward Total Utilitarianism because they show that we cannot escape the Repugnant Conclusion without being forced to accept even more unpalatable conclusions. See also Zuber et al. (2021). Besides money-pump arguments, other arguments for Expected Utility Theory are, for example, the arguments from representation theorems and the long-run argument. See Briggs (2019) and Zynda (2000).

[7] Let $\mathrm{W}(A)$ denote the total quantity of well-being in the state of affairs A and let $\mathrm{w}(S_i)$ denote the well-being of individual S_i. Then,

$$\mathrm{W}(A) = \sum_{i=1}^{\infty} w(S_i).$$

just in case its expected utility is greater.[8] The expected utility of a prospect is calculated by multiplying the utilities of its outcomes by their probabilities, and summing these up. Let EU (X) denote the expected utility of prospect X. Then, more precisely, Expected Utility Theory states the following:

> **Expected Utility Theory:** For all prospects X and Y, X is at least as good as Y if and only if EU(X) is at least as great as EU(Y).[9]

1.2 Boundedness

Standard versions of expected utility maximization require bounded utilities.[10] If utilities are real-valued, then boundedness means the following:

> **Boundedness:** There is some real number M such that for all outcomes x, the utility of x is below M and above $-M$.

In other words, Boundedness rules out arbitrarily and infinitely good outcomes.

Consider for example the von Neumann-Morgenstern Expected Utility Theory.[11] Let XpY be a risky prospect with a p chance of prospect X obtaining and a $1 - p$ chance of prospect Y obtaining (that is, the agent gets probability p of prospect X and probability $1 - p$ of prospect Y). Then, if prospects are compared by their expected utilities, Boundedness follows from the following von Neumann-Morgenstern axiom:

---

[8] In the case of Total Utilitarianism, 'better' is used in an axiological sense; in the case of Expected Utility Theory, 'better' is concerned with instrumental rationality.

[9] Let O be the set of possible outcomes, $p_X(o)$ the probability of outcome o in prospect X and $u(o)$ the utility of o. Then,

$$\text{EU}(X) = \sum_{o \in O} p_X(o)u(o).$$

[10] See for example Kreps (1988, pp. 63–64), Fishburn (1970, pp. 194, 206–207), Hammond (1998, pp. 186–191) and Russell & Isaacs (2021).

[11] The following axioms together entail Expected Utility Theory: Completeness, Transitivity, Independence and Continuity. See von Neumann & Morgenstern (1947), Jensen (1967, pp. 172–182) and Hammond (1998, pp. 152–164).

> **Continuity:** If X is better than Y, which is better than Z, then there are probabilities p and $q \in (0, 1)$ such that XpZ is better than Y and Y is better than XqZ.

For example, suppose a coin is flipped, and an agent gets X with heads and Z with tails. Suppose further that it is possible to alter the bias of the coin. Continuity requires that, with some bias, the agent prefers the coin flip to certainly getting Y, but with some other bias, the agent prefers certainly getting Y to the coin flip.

To see why Continuity implies Boundedness, let's consider the two ways in which Boundedness might be false: arbitrarily and infinitely good prospects. First, Boundedness might be false because there is an infinite sequence of prospects A_1, A_2, A_3, … such that A_2 is at least twice as good as A_1, A_3 is at least twice as good as A_2, and so on, with respect to some baseline. For example, suppose A_1 certainly gives two utilities, A_2 four utilities, A_3 eight utilities, and so on. Next, let A be a mixed prospect that assigns probability $1/2^k$ to prospect A_k (that is, probability $1/2$ to prospect A_1, probability $1/4$ to prospect A_2, probability $1/8$ to prospect A_3, and so on). Then, the expected utility of prospect A is infinite:[12]

$$\text{EU}(A) = 2 \cdot \frac{1}{2} + 4 \cdot \frac{1}{4} + 8 \cdot \frac{1}{8} + ... = 1 + 1 + 1 + ... = \infty.$$

Next, choose some prospects B and C such that B is better than C, and neither has infinite expected utility or disutility. Then, we have that A is better than B (given that A has infinite expected utility), and B is better than C. However, for all non-zero probabilities p, the expected utility of the mixed prospect ApC is infinite. Therefore, ApC is better than B for all non-zero probabilities p. This is a violation of Continuity; there is no probability above zero with which ApC is worse than B.[13]

Secondly, and more generally, Boundedness is false if some prospect A is infinitely better than another (good) prospect B. Similarly as above, this leads to a violation of Continuity because the mixed prospect ApC (where C certainly gives

[12] See Peterson (2022) on the St. Petersburg Paradox.

[13] This is a modified argument from Kreps (1988, pp. 63–64).

nothing) is better than B for all non-zero probabilities p; it is better to get any probability of an infinitely good prospect (and otherwise nothing) than certainty of a finitely good prospect.

To conclude the argument, the two ways in which Boundedness might be false both lead to violations of Continuity. Therefore, it follows from Continuity that Boundedness is true.[14]

2 Bounded Expected Totalism

This section presents Bounded Expected Totalism in more detail.

2.1 Non-linear social transformation function

Let *well-being* refer to how good some outcome is for an individual. And, let *social utility* refer to how good some outcome is overall, from an axiological point of view. Also, let *expected individual utility* represent how good some prospect is for an individual, and let *expected social utility* represent how good some prospect is overall. In the context of Expected Utility Theory, these will be denoted by $\mathrm{EU}_{\mathrm{Ind}}$ and $\mathrm{EU}_{\mathrm{Soc}}$, respectively. In general, *individual betterness* will be used to refer to betterness from an individual's point of view. Similarly, *overall/impersonal betterness* will be used to refer to betterness from a moral point of view.

To combine Total Utilitarianism and Expected Utility Theory, we need a *social transformation function* that takes the total quantity of well-being as input and gives social utilities as output. This transformation function must be non-linear if an infinite or arbitrarily large number of happy individuals might exist, as then the total sum of individuals' well-being might be infinite or arbitrarily large (and similarly

[14]These arguments show that Continuity implies an upper bound on utilities. One can give similar arguments to show that Continuity implies a lower bound on utilities.

for negative well-being).[15,16] But, as Bounded Expected Totalism requires expected social utilities to be bounded, the expected social utilities assigned to prospects that might result in an infinite or arbitrarily large number of happy individuals must be bounded.[17]

One might object that the total quantity of well-being cannot be infinite or arbitrarily large because there is an upper limit to how many individuals might exist. This upper limit might be due to, for example, the Universe being finite. However, this may not be true, so we need a decision theory that can also handle these possibilities.[18] If there is even a tiny probability that an infinite or arbitrarily large number of individuals exist, then the transformation function must be non-linear for utilities to be bounded.

Furthermore, my argument need not rely on infinities. One might have a relatively low upper bound for utility to avoid letting outcomes in which very large numbers of happy individuals exist from dominating the expected utility calculations (and similarly for negative well-being). Doing so requires a non-linear social transformation function if the total quantity of well-being is very large; the non-linear transformation function makes sure that outcomes containing vast quantities of well-being are assigned relatively low utility numbers. Having an upper

[15]Note that the total quantity of well-being is not necessarily infinite if an infinite number of individuals exist. For example, suppose that for each individual $k \in \{1, 2, \ldots\}$, k's well-being measure takes a value in the interval $(0, 2^{-k})$. Then, an infinite number of individuals exist but the total quantity of well-being is bounded. However, this can be ruled out by requiring the individual well-being measures to have the same range.

[16]One might object that Total Utilitarianism is not intended to apply in infinite cases. After all, in infinite cases, the total quantity of well-being is not well-defined. So, one might think that Total Utilitarianism does not make sense if there might be an infinite number of individuals.

[17]Beckstead & Thomas (2023, pp. 6–7) write that Boundedness conflicts with the most natural understanding of utilitarianism as an evaluative theory on which improving n lives by a given amount improves the world by n times as much as improving one life. Similarly, they point out that Total Utilitarianism and its variants put unbounded value on creating good lives.

[18]As Branwen (2009) put it: "Scientists have suggested infinite universes on multiple occasions, and we cannot rule the idea out on any logical ground. Should our theory of rationality stand or fall on what the cosmologists currently think?"

bound for utility is one way to avoid *Probability Fanaticism*:[19]

> **Probability Fanaticism:** For any non-zero probability p, and for any (finitely) good outcome o, there is some great enough outcome O such that probability p of O (and otherwise nothing) is better than certainty of o.

To summarize, social utilities might be bounded if the total quantity of well-being is itself necessarily bounded. However, this is not true. Therefore, Bounded Expected Totalism requires a social transformation function that takes the total quantity of well-being as input and outputs social utilities.

2.2 Defining Bounded Expected Totalism

Suppose that the social transformation function is non-linear. It will also have the following qualities: First, more well-being is always better, so the social transformation function must be strictly increasing with the total quantity of well-being; it must assign greater utilities to outcomes that contain more well-being. Secondly, because utilities are bounded above, similar increases in well-being must (after some point at least) matter less and less. Consequently, the social transformation function must be strictly concave on some subset of its domain. Furthermore, because utilities are also bounded below, similar increases in negative well-being must (after some point at least) matter less and less. Thus, the social transformation function must be strictly convex on some subset of its domain. Lastly, for utilities to be bounded, the social transformation function must be sufficiently concave with positive total well-being and sufficiently convex with negative total well-being; the contribution of additional (positive or negative) well-being to social utility must tend to zero.

Let f be this transformation function. Also, let $p_X(O_i)$ denote the probability of outcome O_i in prospect X and $\mathrm{W}(O_i)$ the total quantity of well-being in O_i.

[19] Wilkinson (2022, p. 449). For discussions related to fanaticism, see Beckstead (2013, ch. 6), Goodsell (2021), Russell & Isaacs (2021), Wilkinson (2022), Beckstead & Thomas (2023) and Russell (2023).

Also, recall that EU_{Soc} denotes expected social utility. Then, we can state Bounded Expected Totalism formally as follows:[20]

> **Bounded Expected Totalism:** For all prospects X and Y, X is at least as good as Y if and only if $\text{EU}_{\text{Soc}}(X)$ is at least as great as $\text{EU}_{\text{Soc}}(Y)$, where
>
> $$\text{EU}_{\text{Soc}}(X) = \sum_{i=1}^{\infty} p_X(O_i) f(\text{W}(O_i)).$$

Bounded Expected Totalism is the view that outcomes are ranked by their total quantities of well-being, and prospects are ranked by expected social utilities, where social utility is some bounded function of the total quantity of well-being. On Bounded Expected Totalism, when calculating the value of a prospect, one first calculates the total quantity of well-being in every possible state of the world. Then, one transforms each state's total quantity of well-being into social utilities. Finally, to get the expected social utility of a prospect, one multiplies the social utility of each state with that state's probability and sums these up.

Next, it will be shown that Bounded Expected Totalism violates Ex Ante Pareto:

> **Ex Ante Pareto:** For all prospects X and Y, if X is at least as good as Y for each individual, and X is better than Y for some, then X is overall better than Y.

[20] This paper discusses what might be called *Ex-Post Bounded Expected Totalism*. However, there is another way Bounded Expected Totalism can deal with risk, which could be called *Ex-Ante Bounded Expected Totalism*. Formally, *Ex-Ante* Bounded Expected Totalism states the following:

> ***Ex-Ante* Bounded Expected Totalism:** For all prospects X and Y, X is at least as good as Y if and only if $\text{EU}_{\text{Soc}}(X)$ is at least as great as $\text{EU}_{\text{Soc}}(Y)$, where
>
> $$\text{EU}_{\text{Soc}}(X) = f\left(\sum_{i=1}^{\infty} p_X(O_i)\text{W}(O_i)\right).$$

Ex-Ante Bounded Expected Totalism violates Continuity. For example, let A be a St. Petersburg-style lottery (with the outcomes being total quantities of well-being), B a prospect that certainly gives a modest good outcome and C a prospect that certainly gives nothing. The expected total well-being of the mixed prospect ApC is infinite for all $p \in (0, 1)$. Thus, the expected social utility of ApC equals the upper bound of utilities, which is greater than the expected social utility of B. So, A is better than B, which is better than C, but ApC is better than B for all $p \in (0, 1)$.

3 The Ex Ante Pareto violations

This section gives two examples to show that Bounded Expected Totalism violates Ex Ante Pareto if social utilities are bounded above and below. These examples show that a violation of Ex Ante Pareto happens regardless of whether individual betterness is risk-neutral, risk-averse or risk-seeking with respect to well-being.

Consider the following prospects:

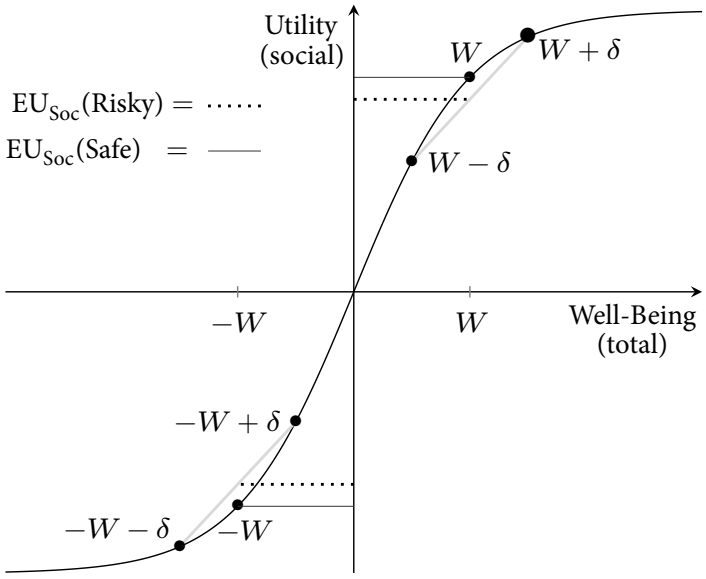
Risky vs. Safe:

Risky Gives a 0.5 probability of δ additional well-being for some individual; otherwise, it decreases their well-being by $-\delta$.

Safe Does not increase or decrease well-being.

If social utilities are bounded above, then (at least at some point) the social transformation function is concave with a positive total quantity of well-being. This means that, at least sometimes, the overall betterness relation is risk-averse with respect to well-being. So, with some positive total quantity of well-being W, Safe is impersonally better than Risky. On the other hand, if social utilities are bounded below, then (at least at some point) the social transformation function is convex with a negative total quantity of well-being. This means that, at least sometimes, the overall betterness relation is risk-seeking with respect to well-being. Thus, with some negative total quantity of well-being $-W$, Risky is impersonally better than Safe. This is illustrated by the following graph:

Risky vs. Safe



So, whether Risky is overall better than Safe (or vice versa) depends on the total quantity of well-being. However, whether Risky is better than (or equally as good or worse than) Safe for some agent Alice does not depend on the total quantity of well-being.

First, suppose that Risky is better than Safe for Alice (and no one else is affected by the choice). Then, Bounded Expected Totalism violates Ex Ante Pareto when the total quantity of well-being in the background population is high (W) because then Safe is better than Risky impersonally. In that case, individual betterness for Alice is risk-seeking with respect to well-being, but overall betterness is risk-averse.

Next, suppose that Safe is better than Risky for Alice (and again equally good for everyone else). Then, Bounded Expected Totalism violates Ex Ante Pareto when the total quantity of well-being in the background population is low ($-W$) because then Risky is better than Safe impersonally. In that case, individual betterness for Alice is risk-averse with respect to well-being, but overall betterness is risk-seeking. So, this case shows that Bounded Expected Totalism violates Ex Ante

Pareto if individual betterness for Alice deviates from risk-neutrality with respect to well-being.[21]

Finally, suppose that Risky and Safe are equally good for Alice. This means that individual betterness for Alice is risk-neutral with respect to well-being. We may change the previous case slightly to show that a violation of Ex Ante Pareto happens again:[22]

The Risk-Neutral Case:

Risky Gives a 0.5 probability of $\delta + \epsilon$ additional well-being for some individual; otherwise, it decreases their well-being by $-\delta$.

Safe Does not increase or decrease well-being.

Given that individual betterness for Alice is risk-neutral with respect to well-being, Risky is better than Safe for Alice as it gives her a higher expectation of well-being (with all positive values of δ and ϵ). However, given that overall betterness is risk-averse when the total quantity of well-being in the background population is high (W), we can find some positive constants δ and ϵ such that Safe is impersonally better than Risky when the total quantity of well-being in the background population is W.[23] Then, Risky is better than Safe for Alice (and equally good for

[21] This case also shows that Bounded Expected Totalism violates Weak Ex Ante Pareto if individual betterness for Alice is risk-neutral with respect to well-being.

[22] Gustafsson (2022*a*) presents a similar case to illustrate that *Ex-Post* Prioritarianism violates Ex Ante Pareto, a fact that goes back at least to Rabinowicz (2002). For an overview of this topic, see Fleurbaey (2018). See also Broome (1991, Ch. 9). Bounded Expected Totalism differs from *Ex-Post* Prioritarianism because it first sums up everyone's well-being and then converts this sum into social utilities. In contrast, the latter view first converts individuals' well-being levels and then sums up the converted well-being levels. However, Bounded Expected Totalism coincides with *Ex-Post* Prioritarianism in one-person cases.

[23] More formally, given that the social transformation is concave at W, there must be some positive constants δ and ϵ such that $f(W) - f(W - \delta) > f(W + \delta + \epsilon) - f(W)$. This is because the smaller benefit (δ) contributes more when added to a population at a lower well-being level than the greater benefit ($\delta + \epsilon$) when added to a population at a higher well-being level. The expected social utility of Safe is $\mathrm{EU}_{\mathrm{Soc}}(\text{Safe}) = f(W)$. And, the expected social utility of Risky is $\mathrm{EU}_{\mathrm{Soc}}(\text{Risky}) = \frac{1}{2} \cdot f(W + \delta + \epsilon) + \frac{1}{2} \cdot f(W - \delta)$. Given that $f(W) - f(W - \delta) > f(W + \delta + \epsilon) - f(W)$, $\mathrm{EU}_{\mathrm{Soc}}(\text{Risky})$ is less than $\mathrm{EU}_{\mathrm{Soc}}(\text{Safe})$.

everybody else) but Safe is better than Risky impersonally. So, we have a violation of Ex Ante Pareto.[24]

To summarize, this section presented two examples to show that Bounded Expected Totalism violates Ex Ante Pareto regardless of whether individual betterness is risk-neutral, risk-averse or risk-seeking with respect to well-being.

4 Harsanyi's social aggregation theorem

This section discusses how the earlier examples relate to a famous result in this area, namely, Harsanyi's social aggregation theorem. Harsanyi's social aggregation theorem shows that if both individual and social betterness relations can be given an expected utility representation, and the overall betterness relation satisfies Ex Ante Pareto, then social utilities are weighted sums of individual utilities.[25]

Let me explain Harsanyi's premises in more detail. The first premise states that each individual's betterness relation obeys the von Neumann-Morgenstern axioms.[26] So, the individual betterness relation can be represented by an expectational utility function. The second premise says that the overall betterness relation obeys the von Neumann-Morgenstern axioms. So, overall betterness can also be represented by an expectational utility function. The third premise is Ex Ante Pareto.[27] The conclusion of Harsanyi's theorem is that social utilities are weighted

[24] This case shows that Bounded Expected Totalism violates Ex Ante Pareto if individual betterness is risk-neutral and utilities are bounded above. One can change the case to show that Ex Ante Pareto is also violated if individual betterness is risk-neutral and utilities are bounded below.

[25] Harsanyi (1955). Harsanyi (1955) uses individual utilities to describe individual preferences. But we may reinterpret them as describing individual betterness instead of individual preferences. See Broome (1991).

[26] Harsanyi (1955) uses Marschak's (1950) versions of the von Neumann & Morgenstern (1947) axioms. Marschak's (1950, p. 117) Postulate II states:

> **Postulate II (Continuity):** If X is better than Y, which is better than Z, then there is a probability $p \in (0, 1)$ such that Y is equally as good as XpZ.

This postulate implies, in a similar way as shown earlier, that utilities must be bounded.

[27] Harsanyi (1955) uses Pareto Indifference in the original formulation of the theorem, while

sums of individual utilities. Thus, overall betterness can be represented as maximizing the expectation of a weighted sum of individual utilities. If, in addition, we assume equal weighting for all individuals, then this theorem shows that the social utility function must be a sum of individual utilities.[28]

Harsanyi's theorem shows, in other words, that if individual and overall betterness relations are represented by expectational utility functions, then in order to satisfy Ex Ante Pareto, the social utility function must be a linear combination of individual utilities. Earlier in this paper, it was shown that Total Utilitarianism combined with bounded Expected Utility Theory violates Ex Ante Pareto. Therefore, if one accepts Bounded Expected Totalism, that premise of Harsanyi's theorem fails. The reason that led to its failure was that a non-linear social transformation function is needed 1.) because the number of individuals might be infinite or arbitrarily large and 2.) to avoid letting outcomes in which very large numbers of happy individuals exist from dominating the expected utility calculations.

In fact, it is unsurprising that one of Harsanyi's premises must be rejected; if the number of individuals might be infinite or arbitrarily large, then social utilities cannot be weighted sums of individual utilities because this could lead to unbounded social utilities.[29,30] Similarly, if one uses a non-linear transformation function to avoid Probability Fanaticism, then social utilities cannot be weighted sums of individual utilities because these sums might exceed the upper bound of utilities. So, given that a bounded expected totalist rejects Harsanyi's conclusion, they cannot accept all his premises.

Harsanyi (1977, p. 65) uses Weak Ex Ante Pareto. Using Weak Ex Ante Pareto instead of Pareto Indifference guarantees that positive individual well-being contributes *non-negatively* to social utilities. Using Ex Ante Pareto instead of Weak Ex Ante Pareto guarantees that positive individual well-being contributes *positively* to social utilities. See Weymark (1994) on Harsanyi's theorem with different Pareto principles.

[28] Broome (1991, §10) argues that Harsanyi's social aggregation theorem, together with Bernoulli's hypothesis, leads to utilitarianism.

[29] See Blackorby et al. (2007) for an extension of Harsanyi's social aggregation theorem to variable populations.

[30] As mentioned earlier, this need not be true. See footnote 15.

This is worrying because Harsanyi's theorem is often considered one of the best arguments for utilitarianism. The conclusion of Harsanyi's theorem is that, for any fixed and finite population, social utility is an affine (or linear) function of total individual utility. However, once we consider the possibility of an infinite or arbitrarily large population (or wish to avoid fanatical prescriptions), we find that social utility must be non-linear if social utilities are bounded and additive with individual utilities. And this leads to violations of Ex Ante Pareto.

These cases might be taken to undermine Boundedness (and Continuity). One might accept, for example, *Unbounded Expected Totalism*, namely, the view that combines Total Utilitarianism and Expected Utility Theory with an unbounded utility function. However, this view cannot be supported by a version of Harsanyi's theorem that relies on the von Neumann-Morgenstern axiomatization of Expected Utility Theory, as this axiomatization has Continuity as one of its axioms. But one might attempt to justify Unbounded Expected Totalism with a Harsanyi-style argument that does not rely on Continuity.[31]

Alternatively, the arguments in this paper might be taken to indirectly support alternatives to Expected Utility Theory and Boundedness, such as discounting small probabilities. Discounting small probabilities also leads to violations of Ex Ante Pareto.[32] But given that both Bounded Expected Totalism and discounting small probabilities violate Ex Ante Pareto, the plausibility of Ex Ante Pareto does not favor the former over the latter.

[31] Fleurbaey (2009) gives such an argument using statewise dominance and anonymity instead of the von Neumann-Morgenstern axioms. Relatedly, McCarthy et al. (2020) show that suitable forms of Pareto and anonymity imply a kind of fixed population expected utilitarianism, but with the individual utilities being vectors (see Theorem 4.10). Furthermore, one can derive Unbounded Expected Totalism with the help of further assumptions, for example, the assumption that there is a constant critical level c such that adding a life at level c is always socially neutral (personal correspondence).

[32] See Kosonen (2021, §5).

5 Conclusion

This paper first discussed how Expected Utility Theory with a bounded utility function can be combined with Total Utilitarianism. Then, it showed that the combination of these views, Bounded Expected Totalism, violates Ex Ante Pareto. Separate examples of Ex Ante Pareto violations were given for risk-neutrality and risk-aversion/risk-seeking. Also, the implications of this case for Harsanyi's social aggregation theorem were discussed. One such implication is that total utilitarians who wish to keep orthodox decision theory can no longer appeal to Harsanyi's argument in support of their theory.

To conclude, combining two standard theories, Total Utilitarianism and Expected Utility Theory with a bounded utility function, results in violations of Ex Ante Pareto: The combination of these views implies that a prospect can be impersonally better than another prospect even though it is not better for anyone who is affected by the choice.

References

Arrhenius, G. (2000), 'An impossibility theorem for welfarist axiologies', *Economics and Philosophy* **16**(2), 247–266.

Beckstead, N. (2013), On the overwhelming importance of shaping the far future, PhD thesis, Rutgers, the State University of New Jersey.

Beckstead, N. & Thomas, T. (2023), 'A paradox for tiny probabilities and enormous values', *Noûs* **00**, 1–25.

Blackorby, C., Bossert, W. & Donaldson, D. (2007), 'Variable-population extensions of social aggregation theorems', *Social Choice and Welfare* **28**(4), 567–589.

Branwen, G. (2009), 'Notes on Pascal's Mugging'.
URL: *https://www.gwern.net/mugging*

Briggs, R. A. (2019), Normative Theories of Rational Choice: Expected Utility, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', Fall 2019 edn, Metaphysics Research Lab, Stanford University.

Broome, J. (1991), *Weighing Goods: Equality, Uncertainty and Time*, Blackwell, Oxford.

Cibinel, P. (2023), 'A dilemma for Nicolausian discounting', *Analysis* **83**(4), 662–672.

Fishburn, P. C. (1970), *Utility Theory for Decision Making*, Wiley, New York.

Fleurbaey, M. (2009), 'Two variants of Harsanyi's aggregation theorem', *Economics Letters* **105**(3), 300–302.

Fleurbaey, M. (2018), 'Welfare economics, risk and uncertainty', *Canadian Journal of Economics* **51**(1), 5–40.

Goodsell, Z. (2021), 'A St Petersburg Paradox for risky welfare aggregation', *Analysis* **81**(3), 420–426.

Gustafsson, J. E. (2022*a*), 'Ex-ante prioritarianism violates sequential ex-ante Pareto', *Utilitas* **34**(2), 167–177.

Gustafsson, J. E. (2022*b*), *Money-Pump Arguments*, Cambridge University Press, Cambridge.

Hájek, A. (2014), 'Unexpected expectations', *Mind* **123**(490), 533–567.

Hammond, P. J. (1998), Objective expected utility: A consequentialist perspective, *in* S. Barberà, P. J. Hammond & C. Seidl, eds, 'Handbook of Utility Theory Volume 1: Principles', Kluwer, Dordrecht, pp. 143–211.

Harsanyi, J. C. (1955), 'Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility', *Journal of Political Economy* **63**(4), 309–321.

Harsanyi, J. C. (1977), *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*, Cambridge University Press, Cambridge.

Isaacs, Y. (2016), 'Probabilities cannot be rationally neglected', *Mind* **125**(499), 759–762.

Jensen, N. E. (1967), 'An introduction to Bernoullian utility theory: I. Utility functions', *The Swedish Journal of Economics* **69**(3), 163–183.

Kosonen, P. (2021), 'Discounting small probabilities solves the Intrapersonal Addition Paradox', *Ethics* **132**(1), 204–217.

Kosonen, P. (2022), Tiny Probabilities of Vast Value, PhD thesis, University of Oxford.

Kreps, D. M. (1988), *Notes on the Theory of Choice*, Westview Press, Boulder.

Lundgren, B. & Stefánsson, H. O. (2020), 'Against the *De Minimis* principle', *Risk Analysis* **40**(5), 908–914.

Marschak, J. (1950), 'Rational behavior, uncertain prospects, and measurable utility', *Econometrica* **18**(2), 111–141.

McCarthy, D., Mikkola, K. & Thomas, T. (2020), 'Utilitarianism with and without expected utility', *Journal of Mathematical Economics* **87**, 77–113.

Monton, B. (2019), 'How to avoid maximizing expected utility', *Philosophers' Imprint* **19**(18), 1–24.

Parfit, D. (1997), 'Equality and priority', *Ratio* **10**(3), 202–221.

Peterson, M. (2022), The St. Petersburg Paradox, *in* E. N. Zalta, ed., 'The Stanford Encyclopedia of Philosophy', Summer 2022 edn, Metaphysics Research Lab, Stanford University.

Rabinowicz, W. (2002), 'Prioritarianism for prospects', *Utilitas* **14**(1), 2–21.

Russell, J. S. (2023), 'On two arguments for fanaticism', *Nous* **00**, 1–31.

Russell, J. S. & Isaacs, Y. (2021), 'Infinite prospects', *Philosophy and Phenomenological Research* **103**(1), 178–198.

Smith, N. J. J. (2014), 'Is evaluative compositionality a requirement of rationality?', *Mind* **123**(490), 457–502.

von Neumann, J. & Morgenstern, O. (1947), *Theory of Games and Economic Behavior*, 2 edn, Princeton University Press, Princeton.

Weymark, J. A. (1994), Harsanyi's social aggregation theorem with alternative Pareto principles, *in* W. Eichhorn, ed., 'Models and Measurement of Welfare and Inequality', Springer, Berlin, Heidelberg, pp. 869–887.

Wilkinson, H. (2022), 'In defence of fanaticism', *Ethics* **132**(2), 445–477.

Zuber, S., Venkatesh, N., Tännsjö, T., Tarsney, C., Stefánsson, H. O., Steele, K., Spears, D., Sebo, J., Pivato, M., Ord, T., Ng, Y.-K., Masny, M., MacAskill, W., Kuruc, K., Hutchinson, M., Gustafsson, J. E., Greaves, H., Forsberg, L., Fleurbaey, M., Coffey, D., Cato, S., Castro, C., Campbell, T., Budolfson, M., Broome, J., Berger, A., Beckstead, N. & Asheim, G. B. (2021), 'What should we agree on about the repugnant conclusion?', *Utilitas* **33**(4), 379–383.

Zynda, L. (2000), 'Representation theorems and realism about degrees of belief', *Philosophy of Science* **67**(1), 45–69.