TRANSFINITELY TRANSITIVE VALUE

By Kacper Kowalczyk[1,2]

This paper develops transfinite extensions of transitivity and acyclicity in the context of population ethics. They are used to argue that it is better to add good lives, worse to add bad lives, and equally good to add neutral lives, where a life's value is understood as personal value. These conclusions rule out a number of theories of population ethics, feed into an argument for the repugnant conclusion, and allow us to reduce different-number comparisons to same-number ones. Challenges to these arguments are addressed, including the issue of comparing existence and non-existence in terms of personal value, the possibility of minimal quanta of time and life, and the meaningfulness of measuring closeness between outcomes with different population sizes. An asymmetry is uncovered between transfinite cycles of worseness and betterness, supporting a version of the weak procreative asymmetry. Transfinite transitivity principles are also favourably compared to the better-known principles of continuity.

Keywords: value theory, transitivity, infinity, population ethics, money pumps, procreative asymmetry.

I. INTRODUCTION

According to

Transitivity of Weak Betterness. For all value-bearers X, Y, Z, if X is at least as good as Y and Y is at least as good as Z, then X is at least as good as Z.

Here is one way to appreciate the meaning of this principle. Take a sequence X_1, X_2, ... such that X_1 is at least as good as X_2, which in turn is at least as good as X_3, which ... and so on. If X comes at the end of such a sequence, transitivity allows us to make a shortcut: we can conclude that X_1 is at least as good as X. Provided that the sequence is finite, that is. *Transfinite* transitivity relaxes this finiteness proviso, allowing us to make infinite as well as finite shortcuts. What this means exactly will be explained in the next section.

Principles of transitivity are powerful, as evidenced by Parfit's mere addition paradox (1987 [1984]: 419–41), Broome's argument against the neutrality intuition (2004: 140–50), and Temkin's (1996) spectra arguments. Rachels (2001)

and Temkin (1996, 2012) take these unexpected results to cast doubt on transitivity itself.

Principles of transfinite transitivity are much more powerful than that. This paper will show how to use them to argue that creating people with good lives is better than not creating them and that creating people with bad lives is worse than not creating them, thus significantly narrowing down the range of possible views in population ethics.

Why believe transfinite transitivity? In my other work, I argue that major arguments in favour of transitivity can also be used to support transfinite transitivity. Its central part concerns transfinite transitivity's role in the theory of rational choice. In an important range of infinite decision problems, transfinite transitivity can play the same role that transitivity plays in finite decision problems: it helps secure the existence of best options, it can be characterized by consistency conditions on permissible choice, and it can be supported by money-pump arguments. This suggests transfinite transitivity truly generalizes transitivity.[1]

Section II explains transfinite transitivity and related principles in more detail. Section III shows how to use them in population ethics. Section IV then takes up some objections and loose ends regarding these arguments. Section V is about how pragmatic (money-pump) arguments bear on the status of these transfinite principles in the context of population ethics. This discussion suggests a novel justification of a version of the so-called procreative asymmetry. Section VI compares transfinite transitivity to a related but better-known principle of continuity, showing how the former is importantly more liberal.

II. PRINCIPLES

Transitivity can be generalized into the transfinite whenever our domain of evaluation is equipped with a notion of *convergence* which is perhaps easiest to explain in terms of *closeness* (or *distance*). We say that the sequence $X_1, X_2, \ldots$ *converges* to X (its *limit*) iff it eventually becomes arbitrarily close to X. For example, the sequence of numbers 1, $\frac{1}{2}$, $\frac{1}{4}$, ... converges to number 0, since the absolute difference between members of that sequence and 0 is eventually arbitrarily small.[2] How to understand closeness in the case of value-bearers will depend on the application at issue, and we will see some examples below.

---

[1] See Kowalczyk (ms). Broome (2004: 50–63) provides a different argument for ordinary transitivity, namely, that it is part of the logic of comparatives. I think a similar argument might well work to support transfinite transitivity, too, the key insight being that the natural ordering of the real line is transfinitely transitive with respect to its usual topology.

[2] More precisely, for any positive number ϵ, all elements sufficiently far in the sequence are within distance ϵ of *x*. This assumes we are working with *metric spaces*, a subset of *topological spaces*,

With convergence in hand, we can build towards *transfinite* transitivity of weak betterness by noting that ordinary transitivity of weak betterness allows us to make shortcuts along finite paths of weak betterness, as follows:[3]

$$X_1 \succcurlyeq X_2 \succcurlyeq X_3 \text{ implies } X_1 \succcurlyeq X_3.$$

$$X_1 \succcurlyeq X_2 \succcurlyeq X_3 \succcurlyeq X_4 \text{ implies } X_1 \succcurlyeq X_4.$$

$$X_1 \succcurlyeq X_2 \succcurlyeq X_3 \succcurlyeq X_4 \succcurlyeq X_5 \text{ implies } X_1 \succcurlyeq X_5.$$

And so on.

Here all $X_1, X_2, \ldots$ are value-bearers. Transfinite transitivity of weak betterness extends this pattern into the transfinite, by means of the notion of convergence: for all value-bearers $X, X_1, X_2, \ldots$,

$$X_1 \succcurlyeq X_2 \succcurlyeq X_3 \succcurlyeq \ldots \to X \text{ implies } X_1 \succcurlyeq X,$$

where '$\to$' indicates convergence.[4] Put differently:

Transfinite Transitivity of Weak Betterness. For all value-bearers $X, X_1, X_2, \ldots$, if the sequence $X_1, X_2, \ldots$ converges to X, and X_1 is at least as good as X_2, X_2 is at least as good as $X_3, \ldots$, then X_1 is at least as good as X.[5]

Analogously, we can define transfinite transitivity of other binary relations such as strict betterness, strict worseness, weak worseness, or equal goodness. Principles from the finite case that are weaker than transitivity can also be generalized into the transfinite. For example, we have

Acyclicity of Weak Betterness. For all value-bearers $X_1, X_2, \ldots, X_n$, if X_1 is at least as good as X_2, X_2 is at least as good as $X_3, \ldots$, and X_{n-1} is at least as good as X_n, then X_n is not better than X_1.[6]

where some distance metric is defined. For more explanation of all mathematical concepts in this paper, see a standard topology textbook such as Mendelson (1962).

[3] Notation: '$\succcurlyeq$' will mean 'at least as good as', '$\succ$' 'better than', '$\prec$' 'worse than', '$\preccurlyeq$' 'at least as bad as', and '$\sim$' will mean 'equally good as'. I will call these relations 'weak betterness', 'strict betterness', 'strict worseness', 'weak worseness', and 'equal goodness', respectively.

[4] This notation is from Bartha, Barker & Hájek (2014) who also introduce the name 'transfinite transitivity' for what I would call 'transfinite acyclicity of strict worseness'. The idea of transfinite transitivity itself is apparently due to Gillies (1959). It was then reinvented by Smith (1974) and discussed by Birchenhall (1977), Mukherji (1977), and later by Carosi & Zaffaroni (1999) and Kukushkin (2008). In philosophy it was reinvented again by Bartha *et al.* (2014). None of these authors have applied transfinite transitivity in ethics, let alone population ethics.

[5] It is easy to see that reflexivity plus transfinite transitivity of weak betterness implies transitivity of weak betterness: if $X \succcurlyeq Y \succcurlyeq Z$, then $X \succcurlyeq Y \succcurlyeq Z \succcurlyeq Z \succcurlyeq \ldots \to Z$, so, by transfinite transitivity, $X \succcurlyeq Z$.

[6] This property is better known as *Suzumura consistency*, due to Suzumura (1976). See also Bossert (2018).

This generalizes to

Transfinite Acyclicity of Weak Betterness. For all value-bearers $X, X_1, X_2, \ldots, X_n$, if the sequence $X_1, X_2, \ldots$ converges to X, and X_1 is at least as good as X_2, X_2 is at least as good as $X_3, \ldots$, then X is not better than X_1.

These definitions work for other binary relations, too. Transfinite transitivity of weak betterness expresses the thought that whenever things are getting worse, in the limit they must be the worst, while transfinite acyclicity of weak betterness expresses the logically weaker thought that whenever things are getting worse, they cannot be better in the limit.

To be clear, these transfinite principles differ from their finite namesakes in that they require some notion of convergence, here understood in terms of closeness (or distance). Whether the principles are appealing or even meaningful depends on that further choice. But, in many cases, we do have an intuitive grasp of convergence.

III. ARGUMENTS

What is the value of adding an extra life? In a sense that is the basic question of population ethics. Transfinite transitivity principles can help us answer it.

III.1 Constant additions

First consider adding a life that is *constantly neutral*, in the sense that it is equally good no matter how long it is. If we graphed the *cumulative value* of such a life against time, that is, the value it would have if it ended at any given time, we would get a flat line. The slope of that graph, the extra cumulative value divided by the time already lived, would always be zero.[7]

To take a specific example, suppose that Zeno is added to some antecedent outcome A with a constantly neutral life n of 80 years. And then his lifespan is repeatedly halved, as in Table 1, where the dash indicates Zeno's non-existence.

Table 1. Shortening Zeno's lifespan

	Zeno	People in A
A_1	Life n	Unaffected
A_2	Life n cut short at 40 years	Unaffected
A_3	Life n cut short at 20 years	Unaffected
...	...	...
A	—	Unaffected

[7] Compare Broome (2004: 68) and Brown (ms). In the end, on page 254, Broome defines 'constantly neutral' in a way which presupposes the coherence of temporal welfare, which my definition does not. And on pages 23–4 he also assumes discrete rather than continuous time. Brown (ms) definition and framework are much closer to mine. His term for 'constantly neutral' is 'flatline'.

Zeno's life n is just as good when it is lived in full as it is when cut short at 40 years, and just as good then as it is when cut short at 20 years, and so on. Hence, each outcome is just as good for Zeno as the next. Since others are unaffected, it is therefore plausible that each outcome is just as good as the next. This follows from

Pareto Indifference. If the same people exist in outcomes X and Y, and everyone is equally well off in X as they are in Y, then X and Y are equally good.

It is also plausible that the sequence of outcomes A_1, A_2, ... converges to outcome A. That is, the sequence of outcomes where Zeno's life is getting shorter and shorter converges to an outcome where Zeno does not exist at all.

Why? First, the region of the spatiotemporal difference between A and members of the sequence is getting smaller and smaller. For i large enough, A_i differs from A only for a fraction of a second. Moreover, the difference does not 'blow up' as it gets smaller and smaller spatiotemporally. The difference between A_i and A is the same as that between A_1 and A, except more localized.

To make the claim of convergence even more plausible we can assume that Zeno's life n is open on the left, so that there is no first moment of time when it is lived. It is therefore like the left-open interval, $(0, 1]$, containing all numbers between zero and one, excluding zero, as opposed to a left-closed interval like $[0, 1]$, which also contains zero. Hence, if Zeno's life n were shortened to zero length, it arguably could not become a point-sized zero-length life, but would have to instead disappear altogether.

We can also think of A_1, A_2, ... as the outcomes of successive finite stages of a *supertask*, a situation where infinitely many actions are performed in finite time, with A being the outcome at the ω-th stage, coming after all the finite stages. For example, compare Zeno's predicament with a version of Benardete's (1964: 259–60) 'paradox of the gods' where a traveller wants to go from point 0 to point 1 but an infinity of impenetrable walls of decreasing thickness is put up in between, with the first wall halfway, at point $\frac{1}{2}$, then halfway between 0 and $\frac{1}{2}$, at point $\frac{1}{4}$, and so on.[8] The traveller therefore has to stay put at point 0. Similarly, if, say, we progressively tweaked Zeno's genes to make him more and more short-lived, his career would never get started.[9]

---

[8] This version is discussed by Peijnenburg & Atkinson (2010). Unlike Benardete's original—discussed by Priest (1999) and Yablo (2000)—this one is unparadoxical, as Peijnenburg and Atkinson also show.

[9] I think both considerations of overall physical similarity and supertasks give us a handle on the notion of convergence. The latter might actually depend on the former. To determine what would happen at the ω-th stage of supertasks, authors such as Allis & Koetsier (1995) and Earman & Norton (1996) appeal to principles of physical continuity, thus presupposing physical convergence. Recent discussion of convergence and closeness in physics can be found in Fletcher (2020). If physics is too contingent for value theory, we might take inspiration from Lewis (1983) and define closeness and convergence in terms of perfectly natural properties instead.

Someone could object that outcomes which differ in the number of people in existence are always some fixed distance apart (and, so, no convergence could obtain between them). But this objection not only goes against the considerations of overall physical similarity and supertasks just mentioned, but also seems to implausibly presuppose that the number of people corresponds to an independent dimension of closeness.

Granting the claim of convergence, Pareto indifference and the assumption that Zeno's life n is constantly neutral imply the following pattern:

$$A_1 \sim A_2 \sim \ldots \to A.$$

Hence, by transfinite transitivity of equal goodness:

$$A_1 \sim A.$$

Since we assumed nothing special about Zeno or about other people in A, we can conclude, in general, that adding *any* constantly neutral and left-open life to *any* outcome is equally good as not adding it.[10]

We can immediately lift the restriction to left-open lives by using ordinary transitivity and Pareto indifference plus the extra assumption that *someone could* have a constantly neutral and left-open life.[11] That extra assumption is plausible enough. If time really is a continuum, why cannot someone's life occupy a temporal region open on the left?[12]

To see how the argument goes from there, consider Table 2.

Table 2. Left-closed additions

	Zeno	People in A
A	—	Unaffected
B	Life n, left-open	Unaffected
C	Life n, left-closed	Unaffected

---

[10] Transfinite *acyclicity* of equal goodness gives a correspondingly weaker claim that the outcome of adding a constantly neutral life is neither better nor worse than the outcome of not adding it.

[11] This makes my argument unlike otherwise similar arguments in Bader (forthcoming, ms-a). Other differences include that Bader does not use constantly neutral lives, nowhere appeals to transfinite transitivity/acyclicity, and mostly works directly with permissible choice rather than value. Nonetheless, this paper would not have existed without the inspiration of Bader's earlier work.

[12] While plausible enough, it is part of an ancient debate about the boundaries of ordinary objects in continuous space; see Varzi (2015). Suffice to say, I have Leonardo da Vinci and Bertrand Bolzano on my side.

Think of outcome A as status quo. Zeno can then be added with a constantly neutral life n which will be left-open in outcome B and left-closed in outcome C.

First, if the argument above goes through for left-open lives, then A and B are equally good. The second step begins by noting that B and C are also equally good for Zeno. His life is longer by an instant in the latter outcome, but that should not matter since it is constantly neutral either way. And since other people are unaffected, it follows from Pareto indifference that B and C are equally good.

A stronger claim seems true as well: if someone's life in two outcomes is identical except that it is left-open in one and left-closed in the other, then both outcomes are equally good for that person. By analogy, recall that, in calculus, the area under a curve between two bounds *a* and *b* is the same whether or not the bounds themselves are included: a strip of no width can add nothing to total area.

Hence, A and B are equally good, and B and C are equally good. So, by transitivity of equal goodness, A and C are equally good. It does not matter whether a constantly neutral life is left-open or not. In general, we get

> **The Constant Addition Principle**. If outcomes X and Y differ only in that there is one extra person in Y living a constantly neutral life, then Y is equally good as X.

We can run direct arguments for similar principles on the negative and the positive sides.[13] For the former, consider adding a life which is *constantly bad* in the sense that it is worse the longer it is. If we graphed its cumulative value against time, we would get a downward-sloping curve. We can then imagine adding Zeno with such a life to outcome A and then repeatedly halving his lifespan. That each outcome in the resulting sequence is better than the last follows from a strengthening of Pareto indifference,

> **Strong Pareto**. If the same people exist in outcomes X and Y, and everyone is at least as well off in X as they are in Y, then X is at least as good as Y, and if, in addition, some are better off in X than they are in Y, then X is better than Y.

Then the relevant claim of convergence and transfinite transitivity of strict worseness imply that adding Zeno with a constantly bad life is worse than not adding him.

On the positive side, the argument is similar. We make Zeno's life *constantly good*, so that it is better the longer it is. It then follows, this time using transfinite

---

[13] This matters since *transfinite* transitivity of *equal goodness* (much like ordinary transitivity of *equal goodness*) is not easily supported by its role in rational choice theory (where strict value relations matter more). Moreover, transitivity of indifference (the preferential counterpart of equal goodness) was historically the most readily rejected form of transitivity, for example, by Armstrong (1948: 3): 'That indifference is not transitive is indisputable, and a world in which it were transitive is indeed unthinkable'. See also Fishburn (1970).

transitivity of strict betterness, that adding Zeno with such a life is better than not adding him. In either case the implicit restriction to left-open lives is easily removed as before.

We therefore obtain *positive* and *negative* versions of the constant addition principle, with 'constantly neutral' and 'equally good as' replaced by 'constantly good' and 'better than', and 'constantly bad' and 'worse than', respectively.

III.2 Mere additions

These principles tell us something about adding lives with a uniform trajectory through time, *constant additions*, as it were. What do they imply about other *mere additions*?

We can use the special status of constantly neutral lives to answer this question. Say we are wondering whether it is better to add Zeno with some life ℓ to outcome A, as in Table 3.

Table 3. Mere additions

	Zeno	People in A
A	—	Unaffected
B	Life n	Unaffected
C	Life ℓ	Unaffected

That is, we are wondering whether C is better than A. To find out, it is enough to compare life ℓ with life n, a constantly neutral one. If ℓ is better than n, then it follows, by strong Pareto, that C is better than B. Since, by the constant addition principle, B is equally good as A, it follows, by ordinary transitivity, that C is better than A. So, in general, it is always better to add a life that is better than some constantly neutral life. We can put this by saying that lives better than some constantly neutral life have *positive contributive value*. Analogously we can also show that lives worse than (just as good as) some constantly neutral life have *negative (zero) contributive value*.

In fact, a comparison with a constantly neutral life is not only sufficient but also necessary to determine a life's contributive value. Suppose, for example, that it is better to add Zeno with ℓ to A. That is, C is better than A. Since, by the constant addition principle, A is equally good as B, it follows by ordinary transitivity that C is better than B. But note that these two differ only in that Zeno has life ℓ in the former and life n in the latter. Since no one else is affected, it is plausible to think that ℓ must therefore be better than n.[14] It would then

---

[14] There are two ways this can be argued. First, if n is at least as good as ℓ, then it follows, by strong Pareto, that B is at least as good as C, contradicting the claim that C is better than B.

follow that if a life has positive contributive value, it must be better than a constantly neutral life. Analogously we can show that lives with negative (zero) contributive value are worse than (just as good as) some constantly neutral life.[15]

If we accept both directions of this argument, we see that constantly neutral lives neatly divide all possible lives in terms of their contributive value. In this way we know the status of all mere additions, not just constant additions.[16]

III.3 Implications for population ethics

Note that, so far, we did not have to assume that welfare can be numerically measured in any fine-grained manner, and we assumed nothing about what it takes for a life to be worth living or worth not living.

By contrast, theories and principles in population ethics typically assume something about both. First, they typically assume that welfare is interpersonally measurable on a ratio scale where, second, zero corresponds to the welfare level of a life which is *neutral* in terms of personal value: on the boundary between worth living and worth not living.

Let's now make these assumptions, too, to see the implications of my arguments for standard claims and theories in population ethics. In this framework, the key question becomes 'What is the welfare level of constantly neutral lives?' Given the argument so far, it is tempting to accept what Brown (ms) would call:

> **The Flatline Analysis of Neutrality**. A life is neutral iff it is just as good as a constantly neutral life.

Brown (ms) is sympathetic to this analysis and does much to defend it. Broome (2004: 68) suggests a similar analysis, too. Accepting it allows us to relate the conclusions reached so far with well-known theses in population ethics.

First, we get the nicely harmonious principle:

> **The Equivalence of Personal and Contributive Value**. The sign of a life's contributive value is equal to the sign of its welfare.[17]

Hence, ℓ is either better than n or incomparable with it. We can rule out the latter possibility by assuming that all lives are comparable in personal value. Secondly, we can use a same-number independence principle (as in Blackorby, Bossert and Donaldson 2005: 159) without assuming comparability.

[15] This claim is compatible with some lives being incomparable with constantly neutral ones, as suggested by Gustafsson (2020). These lives would then have *undefined contributive value*: adding them would be incomparable with not adding them.

[16] The arguments of this subsection and the last also imply that all constantly neutral lives are equally good, worse than constantly good lives, and better than constantly bad lives.

[17] Compare Gustafsson (2020: 87), who traces this sort of principle back to Rabinowicz (2009).

This principle rules out theories for which the two signs can differ.[18] One example is

Average Utilitarianism. Outcome X is at least as good as outcome Y iff average welfare in X is at least as high as average welfare in Y.

This is because if the average welfare of an antecedent population is negative enough, adding an extra life with negative welfare might increase the average.[19] Another example is

Critical-Level Utilitarianism. Outcome X is at least as good as outcome Y iff total welfare in X is at least as high as in Y after positive constant α, the critical-level parameter, was subtracted from each person's welfare in both X and Y.[20]

Hence, a life might have positive welfare but contribute negatively to the value of the outcome if it happens to be below α.[21]

Another upshot is that we can quickly get to

The Repugnant Conclusion. For any outcome in which many people exist, at a high welfare level, there is a better outcome in which many more people exist, at a barely positive welfare level.[22]

To see this, start with some outcome A where many people exist at a high welfare level. Keep adding constantly neutral lives until we get an outcome, say, B, where average welfare is barely positive. By the constant addition principle, each such addition preserves overall value. Hence, by ordinary transitivity, all of them do. So, B is equally good as A. Let Z be an outcome which is like B except that total welfare is slightly higher and equally distributed but where everyone's welfare is still barely positive. That Z is better than B follows from

Non-Anti-Egalitarianism. If the same people exist in outcomes X and Y, and outcome X has higher total welfare, higher average welfare, and is more equal than outcome Y, then X is better than Y.[23]

Hence, Z is better than B, B is equally good as A, so, by transitivity, Z is better than A. This is the repugnant conclusion.[24]

---

[18] If we used transfinite *acyclicity* principles instead of transfinite *transitivity* principles, we would get the weaker claim that the sign of a life's contributive value is *not distinct* from the sign of its welfare, which allows for, say, a positive-welfare life with undefined contributive value.

[19] Compare Parfit's (1987 [1984]: 422) Hell Three.

[20] Compare Broome (2004: 255) and Blackorby *et al.* (2005: 137–8).

[21] This leads to the sadistic conclusion of Arrhenius (2000: 251).

[22] Compare Parfit (1987 [1984]: 388).

[23] See Ng (1989: 238).

[24] This argument is similar to Parfit's (1987 [1984]: 419–41) mere addition paradox. We can see that claims about convergence, transfinite transitivity principles, and Pareto principles can together replace Parfit's mere addition principle, given the flatline analysis of neutrality. This is how the current paper's argument can be read as an impossibility theorem, contributing to the literature exemplified by Ng (1989), Arrhenius (2000), and Blackorby *et al.* (2005: 180–208).

Table 4. Different-number comparisons

	Zeno	Zelda	Xenon
A	10	—	—
B	5	5	5
C	10	$w(\boldsymbol{n})$	$w(\boldsymbol{n})$

In fact, once welfare is interpersonally measurable on a ratio scale, we can draw some important conclusions even without assuming the flatline analysis of neutrality.

The argument for the repugnant conclusion, for example, still works if 'barely positive welfare level' is everywhere replaced by 'welfare level barely above that of constantly neutral lives'. If the original repugnant conclusion is repugnant, it should also be repugnant with that replacement: the Z outcome can differ little from one where everyone's life is on the verge of being always worth ending.

Lastly, given the argument about constant additions, we can reduce comparisons between outcomes with different population sizes to comparisons between outcomes with the same population size. For example, consider Table 4, showing people's welfare levels, with $w(\boldsymbol{n})$ being the welfare level of a constantly neutral life $\boldsymbol{n}$ (not necessarily zero, unless we assume the flatline analysis).

How do A and B compare? From the constant addition principle and Pareto indifference, it follows that A and C are equally good. Hence, by transitivity, A and B compare in the same way that C and B compare. For example, according to

Same-Number Utilitarianism. Outcome X is at least as good as outcome Y iff total welfare in X is at least as high as total welfare in Y, provided that X and Y have the same population size.

Then B is better than C iff

$$3 \times 5 > 10 + 2 \times w(\boldsymbol{n}),$$

which, subtracting $3w(\boldsymbol{n})$ from both sides, becomes

$$3 \times (5 - w(\boldsymbol{n})) > 10 - w(\boldsymbol{n}).$$

So, in general, outcomes compare according to totals of welfare less the welfare level of a constantly neutral life. This is a version of critical-level utilitarianism. Given the flatline analysis of neutrality, $w(\boldsymbol{n})$ becomes zero and we get *total utilitarianism* which simply removes the same-size proviso in same-number utilitarianism. A similar story is true for other same-number principles, such as egalitarianism, prioritarianism, maximin, leximin, and so on.

IV OBJECTIONS AND LOOSE ENDS

I will now consider a number of objections and loose ends that arise regarding the basic argument of the last section, about additions of constantly neutral lives open on the left. First, doesn't the argument overgenerate by implying the contentious claim that a life can be compared with non-existence in terms of personal value? Secondly, is it even possible to make a life arbitrarily short in length? And thirdly, can we really make sense of the idea that a sequence of outcomes with some number of people converges to an outcome with a different number of people?

IV.1 Comparativism

I argued that adding Zeno with a constantly neutral life to any outcome A is just as good as not adding him. Recall Table 5.

Table 5. Shortening Zeno's lifespan

	Zeno	People in A
A_1	Life n	Unaffected
A_2	Life n cut short at 40 years	Unaffected
A_3	Life n cut short at 20 years	Unaffected
...	...	...
A	—	Unaffected

A_1 is therefore equally good as A. But this does not show that A_1 is equally good *for Zeno* as A. Hence, the argument establishes

> **General-Value Comparativism**. Outcomes where one does not exist can be comparable in terms of general value with outcomes where one does exist.

But not necessarily

> **Personal-Value Comparativism**. Outcomes where one does not exist can be comparable in terms of personal value with outcomes where one does exist.[25]

We might think that we could establish the latter by using

> **Transfinite Transitivity of Equal Goodness-For**. For all value-bearers $X, X_1, X_2, \ldots$, the sequence $X_1, X_2, \ldots$ converges to X, and X_1 is equally good for one as X_2, X_2 is equally good for one as X_3, $\ldots$, then X_1 is equally good for one as X.

But opponents of personal-value comparativism will presumably think that A, an outcome where Zeno does not exist, cannot be a bearer of value-for-Zeno.

---

[25] See Broome (1993) for a classic argument against personal-value comparativism, and Bykvist (2007) for a more recent discussion.

And, so, they will think that it does not fall under the antecedent of transfinite transitivity of equal goodness-for. By contrast, A does seem to be a bearer of general value.

The restriction to value-bearers is redundant in the case of ordinary (finite) transitivity where the antecedent already implies that all of the items at issue are value-bearers in the relevant sense. This is not true in the case of *transfinite* transitivity, since the limit outcome is not mentioned in the antecedent at all. There we need to add that restriction explicitly.[26]

IV.2 Divisibility

In the last section, I also assumed that a life can be made arbitrarily short in length. This is to assume something both about the nature of time and the nature of life: that there is no smallest interval of time, and that there is no smallest possible lifespan.

To see that the former is actually needed, suppose that there is a smallest interval of time, say, Δ units. Then we cannot make Zeno's lifespan arbitrarily close to zero since it cannot go down below Δ units. Hence, no sequence of outcomes where Zeno's life gets shorter and shorter converges to an outcome where Zeno does not exist at all.

In response, we can say that continuous time is routinely assumed in physics and that ethics should be no worse off in that respect.[27] So, while the possibility of discrete time is a real challenge, it is not a pressing one.

On the other hand, the possibility that there is a limit on how short a life can be, biologically or psychologically speaking, is less speculative. Yet it was also implicitly assumed away in my argument. Luckily, it can be accommodated, at least to an extent.

To see this, note that there are two ways to read the claim that Δ units, say, is a minimum duration of life:

 (i) Any possible life lasts at least Δ units of time.
(ii) Any possible life lasts more than Δ units of time.

The difference is subtle but important. According to the first, the temporal interval $[0, \Delta]$ is long enough to fit a life: if something starts at time 0 and lasts until time Δ inclusive, it lasts for Δ units in total. But, according to the

---

[26] Even though we lack a non-question-begging argument for personal-value comparativism, we have a non-question-begging argument for the conditional: if personal-value comparativism, then non-existence is just as good for one as existence with a constantly neutral life. Compare Nebel (2019: 325).

[27] Compare, for example, Pruss (2018: 172): 'Standard formulations of major physics theories from Newton onwards either model time with the real numbers or model spacetime as a continuous manifold with local coordinates (...) The continuity involved is essential to the differential equations in which the laws of physics are couched'.

second, that is too short. To be a life, something needs to last longer than any non-life *by a positive margin*.

The last section's argument carries over with small changes on the second reading but fails on the first. To see this, consider Table 6.

Table 6. Shortening Zeno's lifespan to one day

	Zeno	People in A
A_1	Life n, 30,000 days long	Unaffected
A_2	Life n cut short at 15,000.5 days	Unaffected
A_3	Life n cut short at 7,500.75 days	Unaffected
A_4	Life n cut short at 3,750.875 days	Unaffected
. . .	. . .	. . .
A	Life n cut short at one day	Unaffected

Let's assume, for example, that Δ units is one day. A_1 is the outcome of adding Zeno to A with a constantly neutral life n of 30,000 days, which is about 82 years. Then if in some subsequent outcome Zeno's lifespan is T days, we cut it down to $1 + \frac{T-1}{2}$ days in the next. Arguably, the sequence of outcomes $A_1, A_2, \ldots$ converges to A, where Zeno's life is cut short at exactly the one-day mark.

On the first reading, Zeno has a life in all outcomes in the sequence as well as in the limit outcome. Hence, using the strategy of the last section, it follows, at best, that adding a constantly neutral life is just as good as adding a constantly neutral life of minimal duration. But, on the second reading, Zeno does not exist in A. Put differently, 'life n cut short at one day' does not pick out a life but perhaps something which could have been a life if it lasted a little longer. Hence, we can argue as before that adding a constantly neutral life is just as good as not adding it at all.

So, to mount a successful challenge to the argument of the last section, one would have to argue not only that there is a minimum duration of life, but also for a specific reading of 'minimum duration'. But the two readings seem on a par, with the second reading having a slight intuitive edge. So, we again have a real challenge but not a pressing one.

IV.3 Distance metric

But can we even make sense of convergence across population sizes? In some cases this is intuitive enough. But to improve our grip on it, I will now give a toy model of closeness and, so, convergence. It is meant as a consistency check, and perhaps a starting point for more serious models.

In the extant population ethics literature, convergence is sometimes used in the sense of convergence in welfare profile.[28] More precisely, it is often said that the sequence of outcomes $X_1, X_2, \ldots$ converges to outcome X if the welfare profiles of $X_1, X_2, \ldots$ converge to that of X, where a *welfare profile* is an *n*-tuple of real numbers representing the welfare of the n people existing in a given outcome. This assumes, of course, that welfare can be measured on a real-valued scale.

Since a welfare profile is a point in a Euclidean *n*-space, $\mathbb{R}^n$, convergence between welfare profiles can be fixed by the *Euclidean distance*, so that the distance between welfare profiles $(w_1, w_2, \ldots, w_n)$ and $(u_1, u_2, \ldots, u_n)$ is given by

$$\sqrt{(w_1 - u_1)^2 + (w_2 - u_2)^2 + \ldots + (w_n - u_n)^2}.$$

For example, on the real plane ($\mathbb{R}^2$), the Euclidean distance between two points is simply the length of a line segment joining them.

But if two outcomes differ in population size, their welfare profiles are of different dimensions. Hence, their Euclidean distance is undefined, unless we assume the controversial personal-value comparativism.[29] But, for our purposes, there is an easy fix: include information about lifespans in addition to welfare levels.

To see how it goes, think of a life as a point in two-dimensional space, the *x*-axis giving its length, and the *y*-axis giving its overall welfare. The space will look as in Figure 1.

To prop up the last section's argument, we need to make sense of the idea that, as someone's lifespan goes to zero, the whole outcome gets closer to one where they do not exist. But, recall, we only need to make sense of this for lives which are left-open. Hence, we can ignore the possibility that someone's life could correspond to a point directly on the *y*-axis (which is therefore dotted).

So, intuitively, if we make someone's life closer to the *y*-axis, we will move the whole outcome closer to one without them. But to get a well-defined distance, we need to pick a single point on the *y*-axis as reference. Here it is natural to pick the origin, (0, 0).

So, for each person, we can now measure the distance to an outcome without them by simply using the Euclidean distance in a welfare/lifespan space like in Figure 1. For example, the distance between (0, 0) (representing non-existence) and (40, 30) (a life of 40 years with welfare 30) is easily calculated to be 50. We can now aggregate these individual measures into a single measure for whole outcomes.

---

[28] See Blackorby *et al.* (2001) and Broome (2003).

[29] It would also be unwise to *assume* it here, since it could make the last section's argument redundant. See, for example, Holtug (2001: 363–4).
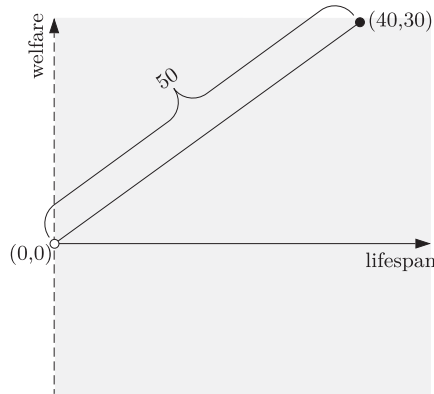
Figure 1. Distance metric.

More precisely, we can define a *welfare-lifespan profile* to be a (countably infinite) list of pairs of welfare levels and *lifespans* for each *possible person*, including possibly the $(0, 0)$ pair (representing non-existence). Examples are $((w_1, l_1), (w_2, l_2), \ldots)$ and $((u_1, k_1), (u_2, k_2), \ldots)$, with $w_1, w_2, \ldots, u_1, u_2, \ldots$ representing welfare levels and $l_1, l_2, \ldots, k_1, k_2, \ldots$ representing lifespans. Then we can define the distance between their corresponding outcomes to be:

$$\sqrt{\begin{array}{c} (w_1 - u_1)^2 + (l_1 - k_1)^2 \\ + (w_2 - u_2)^2 + (l_2 - k_2)^2 \\ + \ldots \end{array}}$$

This new formula can easily be verified to be a distance metric. In cases where population size is finite, it is also always well-defined, since at most a finite number of its terms will be non-zero. In fixed-population fixed-lifespan comparisons, it reduces to the Euclidean distance between welfare profiles. And, lastly, it implies the convergence claims needed for the arguments of the last section.

This toy model also has no untoward substantive implications. For example, even though one's closeness to non-existence is measured from the origin, $(0, 0)$, it does not follow that non-existence is a bearer of personal value in any sense, and a zero-welfare life is not automatically treated as non-existence (since it has positive lifespan).

V PROCREATIVE ASYMMETRY

All arguments in this paper so far were about evaluating outcomes where different numbers of people exist. I will now consider the issue of choosing between such outcomes. In this context, transfinite *acyclicity* becomes more important than transfinite *transitivity*. In particular, transfinite acyclicity of *strict worseness* rather than *strict betterness* turns out to be crucial. This asymmetry between worseness and betterness has significant implications for population ethics and can lead to a form of *weak procreative asymmetry*, according to which, other things equal, there is more reason against adding bad lives than against failing to add good ones.[30]

V.1 Asymmetry of transfinite cycles

Since transfinite transitivity/acyclicity are infinitary principles, to see how they bear on permissible choice, we have to consider choices from infinite sets of outcomes. But infinite sets of outcomes are notoriously problematic.[31] Suppose, for example, the more money the better and consider {$1, $2, $3, . . .}, the set of natural-valued dollar amounts, or [$0, $100), the set of real-valued dollar amounts greater than $0 but strictly less than $100. In either case no outcome is *maximal* in the sense of not being worse than any available outcome. A fortiori, no outcome is *best* in the sense of being at least as good as any available. But it is often assumed that, setting deontological considerations aside, one is permitted to choose all and only maximal outcomes. So, it is puzzling what one is to choose in these cases.

Here the puzzlement arguably arises because the set of available outcomes, *S*, fails to have the property of

> **Compactness**. Every sequence drawn from *S* has a subsequence which converges to something in *S*.

In our examples, the sequence $1, $2, $3, . . . diverges to infinity, so has no convergent subsequence, and even though the sequence $0, $50, $75, $87.5, . . . converges to $100, $100 itself is not an available outcome. Since many of the paradigmatically problematic infinite decision puzzles have to do with non-compactness, it is reasonable to expect compact sets of outcomes to be unproblematic when it comes to the existence of maximal (generally: choiceworthy) outcomes.

---

[30] The procreative asymmetry first appears in Narveson (1973), the label itself being due to McMahan (1981). In its stronger form it says that, other things equal, there is *some* reason against adding bad lives but *no* reason against failing to add good ones. A version of the weak procreative asymmetry is discussed in McMahan (2009).

[31] See Pollock (1983), Slote (1989: 47–81), Sorensen (1994), and Arntzenius, Elga, and Hawthorne (2004).

Hence, it is important news that transfinite cycles of *strict worseness* rule out the existence of maximal elements in compact sets. By contrast, transfinite cycles of *strict betterness* do not need to do so.

For example, the former sort of cycle can arise if adding constantly bad lives is sometimes better than not adding them (a potential implication of average utilitarianism given the flatline analysis of neutrality). To see this, suppose we add Zeno with a constantly bad life $\mathcal{b}$ to some antecedent outcome A and then progressively halve his lifespan.

Table 7. Transfinite cycle of strict worseness

	Zeno	People in A
A_1	Life $\mathcal{b}$	Unaffected
A_2	Life $\mathcal{b}$ cut short at 40 years	Unaffected
A_3	Life $\mathcal{b}$ cut short at 20 years	Unaffected
...	...	...
A	—	Unaffected

If adding Zeno with $\mathcal{b}$ is better than not adding him, then, with strong Pareto, we get the following transfinite cycle of strict worseness:

$$A \prec A_1 \prec A_2 \prec \ldots \rightarrow A.$$

Let $S = \{A, A_1, A_2, \ldots\}$. No outcome is maximal in S, as no outcome of the form A_i can be maximal, since it is worse than A_{i+1}, and A cannot be maximal either, since it is worse than A_1. Yet S is easily verified to be compact.

A mirror image of this situation can arise if adding constantly good lives is sometimes worse than not adding them (a potential implication of critical-level utilitarianism given the flatline analysis of neutrality). In Table 8, $\mathcal{g}$ is some constantly good life.

Table 8. Transfinite cycle of strict betterness

	Zeno	People in A
B_1	Life $\mathcal{g}$	Unaffected
B_2	Life $\mathcal{g}$ cut short at 40 years	Unaffected
B_3	Life $\mathcal{g}$ cut short at 20 years	Unaffected
...	...	...
B	—	Unaffected

If adding Zeno with $\mathcal{g}$ is worse than not adding him, then, with strong Pareto, we get the following transfinite cycle of strict betterness:

$$B \succ B_1 \succ B_2 \succ \ldots \rightarrow B.$$

Let $S = \{B, B_1, B_2, \ldots\}$. If we assume ordinary (finite) transitivity, S does have a maximal, even best, outcome, namely, B. Hence, we see that the two sorts of transfinite cycles (and the theories that can generate them) have different implications for the possibility of maximizing choice, even from compact sets.[32]

V.2 Asymmetry of pumpability

This difference maps on to whether the two sorts of transfinite cycles make one liable to money pumps. We say that an agent is liable to *forcing money pumps* if in some possible situation they *have to* make a series of actions so that they end up paying some cost even though an otherwise identical costless alternative could also be obtained, and that they are liable to *non-forcing money pumps* if in some possible situation they *might* but perhaps do not *have to* make such a series of actions.[33]

We can leave the nature of this cost open. It need not be monetary but can instead be denominated in some other relevant currency. And it need not fall on the agent alone but can instead be spread to some or all of the people involved. The basic idea behind money-pump arguments is that, whatever the cost is, one's theory would not tell one to pay it in advance. Hence, if, together with the structure of one's decision problem, following one's theory means that one ends up paying anyway, then the theory is self-defeating in an objectionable manner.[34]

Let's start with finite cycles. Suppose that A is worse than B, which is worse than C, which is worse than A, and consider the decision tree in Figure 2 where squares represent the agent's choices, leading to further choices or to final outcomes.[35]

Here we offer the agent a choice from A, B, C, and, whatever they choose, we give them a chance to switch to the next-better outcome for a small payment. The claim is that the agent, if guided by value, will end up paying no matter what (as indicated by the arrows). Why?

First, the agent has to choose B over A. This follows if the agent is guided by value in a minimal way, namely, in pairwise choices. But then, since B is

---

[32] Critical-band utilitarianism (as in Blackorby *et al.* 2005: 219–22) implies that adding a good life might be incomparable with not adding it, while some person-affecting views (as in Heyd 1988 and Bader ms-b) imply this about bad lives, too. These views do not lead to transfinite cycles, but at worst to transfinite intransitivities. They therefore avoid the most worrisome problems of average and critical-level utilitarianism discussed in this section.

[33] The distinction appears in Gustafsson & Espinoza (2010).

[34] Liability to money pumps is therefore not merely a sign of imprudence. But this is not the place for a full defence of money-pump arguments in ethics. For a recent sympathetic discussion, see Gustafsson (2015).

[35] This decision tree is due to Cantwell (2003: 389) and also employed by Gustafsson (2015: 1596–97). Money pumps first appear in Davidson, Mckinsey, and Suppes (1955).
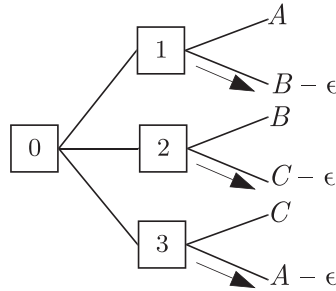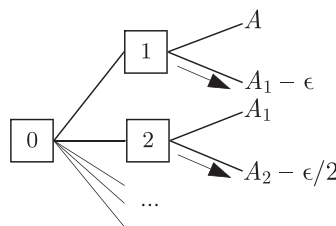
Figure 2. Finite money pump.

strictly better than A, the agent should choose B even at some small cost ϵ. Similarly for C and B, and A and C. Hence, they have to go down at all nodes following the initial one.

Secondly, the agent is permitted to do something at the initial node. From that node's vantage point there are only finitely many outcomes that can be obtained. And, the argument goes, finite sets of outcomes never present moral dilemmas. This is plausible so long as deontological considerations are set aside.[36] So, the agent is permitted to embark on a plan to get one of the obtainable outcomes.

The agent is therefore permitted to go somewhere from the initial node but then they have to go down. Hence, in two moves, the agent is required to pay for something they could have had for free. This is a forcing money pump. Since this is objectionably self-defeating, we can conclude that strict worseness cannot be cyclic after all.[37]

Now consider an infinite version of this pump, as in Figure 3, directed against the transfinite cycle of strict worseness from Table 7.

Figure 3. Transfinite money pump.

[36] Compare Kreps's (2013: 3) assumption of finite non-emptiness in rational choice theory.
[37] There are ways to get around this pump, for example, by adopting *resolute choice*. See Machina (1989) and McClennen (1990). I lack space to defend Cantwell's pump against this possibility.

Here we offer the agent a choice between outcomes A, A_1, A_2, ..., and, whatever they choose, we give them a chance to switch to the next-better outcome for a small payment.[38] The claim is, again, that the agent (an average utilitarian, say), if guided by value, will end up paying no matter what.

The first step is like before (this time backed up by strong Pareto). The second step appeals not to the claim that finite sets never present moral dilemmas, but instead to the claim that *compact* sets do not. While choice from infinite sets of outcomes might be hard, we saw above that many paradigmatic infinite decision puzzles are due to non-compactness. Hence, if we demand choice to be possible from finite sets, we should be happy to demand it from infinite compact sets.[39] In the current example, it also does not *seem* like the agent's initial choice is a moral dilemma: they are effectively asked whether to add Zeno with a constantly bad life and, if so, how long to make it.

Hence, in two moves, the agent is again required to pay for something they could have had for free. This is a forcing money pump. If we should reject cycles of strict worseness in response to the original pump, we should reject *transfinite* cycles of strict worseness here, too.

By contrast, the transfinite cycle of strict betterness from Table 8 does not lead to trouble in the same sort of pump. To see this, consider Figure 4.
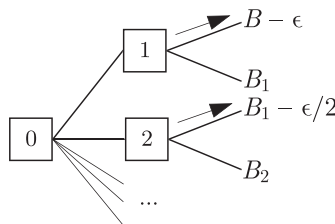
Figure 4. Failed transfinite money pump.

In this case, the agent should arguably end up paying ϵ for B, as B is best in $\{B, B_1, B_2, ...\}$. Since B is not available for free anyway, the agent does not end up paying for something they could have had for free. And if B were also freely available, say, directly from the initial node, the agent should arguably go for that free B.[40]

---

[38] The side payments converge to zero to ensure compactness of the set of obtainable outcomes. It is not necessary that they decrease *geometrically*.

[39] Compactness is routinely assumed in economics in this context, see Kreps (2013: 1–29). And, in mathematics, compactness is seen as a generalization of finiteness, as per Hermann Weyl's apocryphal gloss on one of its many equivalent formulations: 'if a city is compact, it can be guarded by a finite number of arbitrarily near-sighted policemen'; see Hewitt (1960).

[40] This does not, strictly speaking, show that *no* pump is available against a transfinite cycle of strict betterness, although I think that stronger claim is true as well.

V.3 Asymmetry of procreation

What makes one liable to money pumps in this context is not the transfinite cycle of strict worseness itself but rather the transfinitely cyclic pattern of pairwise choices it generates. And that pattern can arise independently if, for example, adding bad lives is sometimes permissible. By contrast, it does not immediately arise if not adding good lives is sometimes permissible.

To get the argument going, we need one extra plausible assumption, namely, that if adding a *bad* life is permissible, then adding an equally good but *constantly bad* life is just as permissible. Without loss of generality, then, suppose that it is permissible to add Zeno with constantly bad life *ɞ* to some antecedent outcome *A*, as in Table 7.

Now consider the corresponding decision tree in Figure 3 again. The claim is that if the agent is allowed to add Zeno with *ɞ* to *A*, then they *might* end up paying no matter what. (If they are *required* to add him, then they *have to* pay.)

The first step is that the agent may go down at all nodes following the initial node. First, the agent should go down at nodes 2, 3, 4, . . . because that is at least as good for all and better for some. This can be backed up by strong Pareto, this time deontically construed. Put differently, it is worth paying something to make Zeno's miserable life twice as short as it would otherwise be. And, second, it arguably follows from our hypothesis that it is permissible to add Zeno with *ɞ* to *A* that the agent *may* go down at node 1 as well.[41]

The next step is the same as before: either compact sets present no moral dilemmas or this particular choice is not a moral dilemma.

Hence, the agent is permitted (at worst, required) to end up paying for something they could have had for free. This is a money pump (at worst, a forcing one).

By contrast, the permission not to add a good life does not land our agent in trouble in the same way. This can be seen by considering the decision tree in Figure 4 again, on the supposition that the agent is permitted not to add Zeno with constantly good life *ɠ*. There the agent is arguably still be permitted to end up with $B - \epsilon$, hence avoiding a money pump.

If arguments in the first part of this paper, based on principles of transfinite transitivity, are sound, then adding a bad life and not adding a good life are both suboptimal: they make things worse. What we can conclude from this section, however, is that there is a special reason against adding bad lives which does not speak against not adding good lives. That is, the permission to do the former can make one liable to money pumps in a way in which

---

[41] If, on balance, one's reasons favour adding Zeno, then that will be true if adding Zeno comes at a small cost. If, on balance, one's reasons favour adding and not adding equally, then we can arrange things so that one's reasons favour adding Zeno more (by making his life slightly better, say). And if the balance of reasons fails to decide, then, plausibly, making Zeno's creation slightly costlier will not change that.

the permission to do the latter does not. Hence, we get a form of the weak procreative asymmetry: there is more reason against adding bad lives than against failing to add good lives, at least other things being equal. Importantly, this asymmetry at the level of choice can be combined with a fully symmetric picture at the level of value.[42]

VI CONTINUITY

Transfinite transitivity and acyclicity principles are *topological principles* in the sense that they make an essential reference to the *topology* of the domain of evaluation such as facts about convergence and closeness. To forestall confusion, it is important to distinguish them from some better-known topological principles, namely, continuity principles. This comparison also serves to underscore the former's appeal.[43]

I will focus on

Transfinite Transitivity of Weak Worseness. For all value-bearers $X, X_1, X_2, \ldots$, if the sequence $X_1, X_2, \ldots$ converges to X, and X_1 is at least as bad as X_2, X_2 is at least as bad as $X_3, \ldots$, then X_1 is at least as bad as X.

Put differently: for all value-bearers $X, X_1, X_2, \ldots$,

$$X_1, X_2, \ldots \to X \text{ and } X_i \preccurlyeq X_{i+1}, \text{ for all } i, \text{ implies } X_1 \preccurlyeq X.$$

Contrast this with

Continuity of Weak Worseness. For all value-bearers $Y, X, X_1, X_2, \ldots$, if the sequence $X_1, X_2, \ldots$ converges to X, and Y is at least as bad as X_1, Y is at least as bad as $X_2, \ldots$, then Y is at least as bad as X.

Put differently: for all value-bearers $Y, X, X_1, X_2, \ldots$,

$$X_1, X_2, \ldots \to X \text{ and } Y \preccurlyeq X_i, \text{ for all } i, \text{ imply } Y \preccurlyeq X.^{44}$$

The idea is that value comparisons (here: weak worseness) are preserved in the limit.

Continuity and transfinite transitivity differ in two ways. First, the $X_1, X_2, \ldots$ sequence has to be $\preccurlyeq$-ordered for transfinite transitivity but not

---

[42] This account of the asymmetry relies on an intrinsic asymmetry between infinite chains of worseness and betterness. It is therefore not ad hoc, unlike many accounts surveyed in Roberts (2011). It also does not rely on substantive deontological resources. It is close in spirit, and inspired by, the account in Bader (forthcoming), although Bader does not appeal to transfinite cycles nor money pumps but instead to principles of universalizability.

[43] Another principle of this sort is *hypersensitivity avoidance*, discussed in Pummer (2019).

[44] In metric spaces, this is the same as the set $\{Z : Y \preccurlyeq Z\}$ being closed for all value-bearers Y. See, for example, Kreps (2013: 13–4) in microeconomics, and Blackorby *et al.* (2005: 92) in population ethics.

for continuity. Secondly, in the case of transfinite transitivity but not continuity, the Y against which members of the sequence are compared has to be the sequence's first member. So, it is easy to see that, given ordinary transitivity, continuity implies transfinite transitivity.[45] Yet the converse fails in an interesting way, since transfinite transitivity, unlike continuity, can accommodate lexical value relations.[46]

To see this, let's use Thomas's (2018: 813) toy example of *total lexical utilitarianism*. To introduce it, first assume that only two things matter for personal value: *love* and *money*. We will think of them as real-valued quantities. Then according to

> **Total Lexical Utilitarianism**. Outcome X is at least as good as outcome Y iff total love in X is greater than total love in Y, or they are equal and total money in X is at least as great as total money in Y.

So, each life and each outcome maps to a point of a two-dimensional space, with total quantities of love on the *x*-axis and of money on the *y*-axis. The horizontal dimension is lexically superior. Since we are in a two-dimensional space, let's assume that the relevant notion of convergence between outcomes is fixed by the Euclidean distance.

It is easy to see that all this implies that Thomas's toy theory violates continuity of weak worseness, as in Table 9.[47]

Table 9. Discontinuity

	Zeno (love, money)	Other people
A_1	(80, 0)	Unaffected
A_2	(40, 0)	Unaffected
A_3	(20, 0)	Unaffected
. . .	. . .	. . .
A	(0, 0)	Unaffected
B	(0,100)	Unaffected

[45] Compare Proposition 3.6 in Smith (1974). Bartha *et al.* (2014: 641–2, 657–8) also discuss the connection between principles they call 'transfinite transitivity' and 'continuity'. But by 'continuity' they mean continuity of the *numerical representation* of the underlying binary relation. As Luce & Suppes (1965: 265) indicate, continuity of the binary relation is the more fundamental notion.

[46] On lexicality, see Debreu (1954). Transfinite transitivity can also accommodate some incomplete value relations that continuity cannot. On incompleteness and continuity, see Aumann (1962: 450–3).

[47] Thomas (2018: fn. 18) says otherwise because he assumes the discrete topology rather than the more natural Euclidean one.

In this example, Zeno's money-deprived life is diminishing in terms of love over the sequence A_1, A_2, ... But so long as Zeno's life holds some love, each of these outcomes is better than B, where he has no love but some money. Yet the sequence A_1, A_2, ... converges to A, where Zeno has no love and no money, which is therefore worse than B.

This is a violation of continuity of weak worseness but not transfinite transitivity of weak worseness. For the latter we would need not only the fact that B is worse than A_1 without being at least as bad as A, but also that all outcomes in the sequence A_1, A_2, ... are at least as bad as the next, which we do not have in this case.

Thus, while proponents of lexical views in population ethics have reasons to be wary of topological principles like continuity, they have no immediate reason to be wary of transfinite transitivity principles. Therefore, framing this paper in terms of transfinite transitivity has a real advantage, besides pumping transitivity-friendly intuitions, even though, logically speaking, much of it would work if we used continuity principles instead.[48]

VII CONCLUSION

This paper introduced and explained infinite extensions of transitivity and acyclicity in the context of population ethics. It showed how to use these principles (given a suitable notion of convergence or closeness) to establish that adding constantly neutral lives is just as good as not adding them, adding constantly good lives is better, and adding constantly bad lives is worse. These arguments have important implications for population ethics: ruling out a number of theories, feeding into arguments for the repugnant conclusion, and reducing different-number comparisons to same-number comparisons. It also addressed a number of challenges to these arguments: the issue of personal-value comparativism, the possibility of minimal quanta of time and life, and provided a toy model of convergence across population sizes. The paper also uncovered an asymmetry between infinite chains of betterness and worseness which leads to an asymmetry in liability to money pumps and to an asymmetry in the ethics of procreation. Transfinite transitivity and acyclicity were also argued to be more appealing than standard topological principles of continuity.[49]

REFERENCES

Allis, V. and Koetsier, T. (1995) 'On Some Paradoxes of the Infinite II', *The British Journal for the Philosophy of Science*, 46: 235–47.

Armstrong, W. E. (1948) 'Uncertainty and the Utility Function', *The Economic Journal*, 58: 1–10.

Arntzenius, F., Elga, A. and Hawthorne, J. (2004) 'Bayesianism, Infinite Decisions, and Binding', *Mind*, 113: 251–83.

Arrhenius, G. (2000) 'An Impossibility Theorem for Welfarist Axiologies', *Economics and Philosophy*, 16: 247–66.

Aumann, R. J. (1962) 'Utility Theory Without the Completeness Axiom', *Econometrica*, 30: 445–62.

Bader, R. (forthcoming) 'The Asymmetry' , in J. McMahan, T. Campbell, K. Ramakrishnan and J. Goodrich (eds.) *Ethics and Existence: The Legacy of Derek Parfit*. Oxford: OUP.

——— (ms-a) 'Axiological Zeno Problems'. Unpublished manuscript.

——— (ms-b) 'Person-affecting Population Ethics'. Unpublished manuscript.

Bartha, P., Barker, J. and Hájek, A. (2014) 'Satan, Saint Peter and Saint Petersburg', *Synthese*, 191: 629–60.

Benardete, J. A. (1964) *Infinity: An Essay in Metaphysics*. Oxford: Clarendon Press.

Birchenhall, C. R. (1977) 'Conditions for the Existence of Maximal Elements in Compact Sets', *Journal of Economic Theory*, 16: 111–5.

Blackorby, C., Bossert, W. and Donaldson, D. (2001) 'Population Ethics and the Existence of Value Functions', *Journal of Public Economics*, 82: 301–8.

——— (2005) *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: CUP.

Bossert, W. (2018) 'Suzumura-Consistent Relations: An Overview', *International Journal of Economic Theory*, 14: 21–34.

Broome, J. (1993) 'Goodness Is Reducible to Betterness: The Evil of Death is the Value of Life', in P. Koslowski and Y. Shionoya (eds.) *The Good and the Economical*, 70–84. Berlin and Heidelberg: Springer.

——— (2003) 'Representing an Ordering When the Population Varies', *Social Choice and Welfare*, 20: 243–6.

——— (2004) *Weighing Lives*. Oxford: OUP.

Brown, C. (ms). 'Better Than Nothing'. Unpublished manuscript.

Bykvist, K. (2007) 'The Benefits of Coming into Existence', *Philosophical Studies*, 135: 335–62.

Cantwell, J. (2003) 'On the Foundations of Pragmatic Arguments', *Journal of Philosophy*, 100: 383–402.

Carosi, L. and Zaffaroni, A. (1999) 'On the Existence of Maximal Elements for Partial Preorders', *Journal of Information and Optimization Sciences*, 20: 271–86.

Davidson, D., Mckinsey, J. C. C. and Suppes, P. (1955) 'Outlines of a Formal Theory of Value, I', *Philosophy of Science*, 22: 140–60.

Debreu, G. (1954) 'Representation of a Preference Ordering by a Numerical Function', in R. M. Thrall, C. H. Coombs and R. L. Davis (eds.) *Decision Processes*, 159–66. New York: Wiley.

Earman, J. and Norton, J. (1996) 'Infinite Pains: The Trouble with Supertasks', in A. Morton and S. P. Stich (eds.) *Benacerraf and His Critics*, 231–61. Oxford: Blackwell.

Fishburn, P. C. (1970) 'Intransitive Indifference in Preference Theory: A Survey', *Operations Research*, 18: 207–28.

Fletcher, S. C. (2020) 'The Principle of Stability', *Philosophers' Imprint*, 20: 1–22.

Gillies, D. B. (1959) 'Solutions to General Zero-Sum Games', in A. Tucker and R. Luce (eds.) *Contributions to the Theory of Games IV*, 47–85. Princeton: Princeton University Press.

Gustafsson, J. E. (2015) 'Sequential Dominance and the Anti-Aggregation Principle', *Philosophical Studies*, 172: 1593–601.

——— (2020) 'Population Axiology and the Possibility of a Fourth Category of Absolute Value', *Economics and Philosophy*, 36: 81–110.

———, and Espinoza, N. (2010) 'Conflicting Reasons in the Small-Improvement Argument', *Philosophical Quarterly*, 60: 754–63.

Hewitt, E. (1960) 'The Role of Compactness in Analysis', *The American Mathematical Monthly*, 67: 499–516.

Heyd, D. (1988) 'Procreation and Value Can Ethics Deal with Futurity Problems?', *Philosophia*, 18: 151–70.

Holtug, N. (2001) 'On the Value of Coming into Existence', *Journal of Ethics*, 5: 361–84.

Kowalczyk, K. (ms) 'Transfinite Transitivity and Rational Choice'. Unpublished manuscript.

Kreps, D. M. (2013) *Microeconomic Foundations*. Princeton: Princeton University Press.

Kukushkin, N. S. (2008) 'Maximizing an Interval Order on Compact Subsets of Its Domain', *Mathematical Social Sciences*, 56: 195–206.

Lewis, D. (1983) 'New Work for a Theory of Universals', *Australasian Journal of Philosophy*, 61: 343–77.

Luce, D. R. and Suppes, P. (1965) 'Preference, Utility, and Subjective Probability', in D. R. Luce, R. R. Bush and E. Galanter (eds.) *Handbook of Mathematical Psychology*, 250–410. New York: Wiley.

Machina, M. J. (1989) 'Dynamic Consistency and Non-Expected Utility Models of Choice Under Uncertainty', *Journal of Economic Literature*, 27: 1622–68.

McClennen, E. F. (1990) *Rationality and Dynamic Choice: Foundational Explorations*. Cambridge: CUP.

McMahan, J. (1981) 'Problems of Population Theory', *Ethics*, 92: 96–127.

——— (2009) 'Asymmetries in the Morality of Causing People to Exist', in M. A. Roberts and D. T. Wasserman (eds.) *Harming Future Persons: Ethics, Genetics and the Nonidentity Problem*, 49–68. Dordrecht, Netherlands: Springer.

Mendelson, B. (1962) *Introduction to Topology*. Boston: Allyn and Bacon.

Mukherji, A. (1977) 'The Existence of Choice Functions', *Econometrica*, 45: 889–94.

Narveson, J. (1973) 'Moral Problems of Population', *Monist*, 57: 62–86.

Nebel, J. M. (2019) 'An Intrapersonal Addition Paradox', *Ethics*, 129: 309–43.

Ng, Y. (1989) 'What Should We Do about Future Generations?', *Economics and Philosophy*, 5: 235–53.

Parfit, D. (1987 [1984]) *Reasons and Persons*. Reprinted with corrections to the 1984 edition. Oxford: OUP.

Peijnenburg, J. and Atkinson, D. (2010) 'Lamps, Cubes, Balls and Walls: Zeno Problems and Solutions', *Philosophical Studies*, 150: 49–59.

Pollock, J. L. (1983) 'How do you maximize expectation value?', *Noûs*, 17: 409–21.

Priest, G. (1999) 'On a Version of One of Zeno's Paradoxes', *Analysis*, 59: 1–2.

Pruss, A. R. (2018) *Infinity, Causation, and Paradox*. Oxford: OUP.

Pummer, T. (2019) 'The Worseness of Nonexistence', in E. Gamlund and C. T. Solberg (eds.) *Saving People from the Harm of Death*, 215–28. New York: Oxford University Press.

Rabinowicz, W. (2009) 'Broome and the Intuition of Neutrality', *Philosophical Issues*, 19: 389–411.

Rachels, S. (2001) 'A Set of Solutions to Parfit's Problems', *Noûs*, 35: 214–38.

Roberts, M. A. (2011) 'An Asymmetry in the Ethics of Procreation', *Philosophy Compass*, 6: 765–76.

Slote, M. (1989) *Beyond Optimizing: A Study of Rational Choice*. Cambridge, MA: Harvard University Press.

Smith, T. (1974) 'On the Existence of Most-Preferred Alternatives', *International Economic Review*, 15: 184–94.

Sorensen, R. (1994) 'Infinite Decision Theory', in J. Jordan (ed.) *Gambling on God: Essays on Pascal's Wager*, 139–59. Lanham, MD: Rowman and Littlefield.

Suzumura, K. (1976) 'Remarks on the Theory of Collective Choice', *Economica*, 43: 381–90.

Temkin, L. S. (1996) 'A Continuum Argument for Intransitivity', *Philosophy and Public Affairs*, 25: 175–210.

——— (2012) *Rethinking the Good*. Oxford: OUP.

Thomas, T. (2018) 'Some Possibilities in Population Axiology', *Mind*, 127: 807–32.

Varzi, A. (2015) 'Boundary', in E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy* (winter 2015 edition). <https://plato.stanford.edu/entries/boundary/> accessed 11 January 2021.

Yablo, S. (2000) 'A Reply to New Zeno', *Analysis*, 60: 148–51.

[1] *St Anne's College, Oxford, UK*

[2] *University College London, UK*