Economics & Philosophy (2024), 1–22
doi:10.1017/S0266267124000099

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Non-Archimedean population axiologies

Calvin Baker

Department of Philosophy, Princeton University, Princeton, NJ, USA
Email: calvin.baker@princeton.edu

(Received 30 March 2023; revised 01 December 2023; accepted 17 December 2023)

Abstract

Non-Archimedean population axiologies – also known as lexical views – claim (i) that a sufficient number of lives at a very high positive welfare level would be better than any number of lives at a very low positive welfare level and/or (ii) that a sufficient number of lives at a very low negative welfare level would be worse than any number of lives at a very high negative welfare level. Such axiologies are popular because they can avoid the (Negative) Repugnant Conclusion and satisfy the adequacy conditions given in the central impossibility result in population axiology due to Gustaf Arrhenius. I provide a novel argument against them which appeals to the way that good and bad lives can intuitively outweigh one other.

Keywords: Population axiology; non-Archimedeanism; lexical views; value superiority; (Negative) Repugnant Conclusion

1. Introduction

By a 'heavenly life' I mean a life at a very high positive welfare level. By a 'barely good life' I mean a life at a very low positive welfare level. This allows us to state the

Repugnant Conclusion.

For any number of heavenly lives, there is a number of barely good lives that would be better.[1]

Similarly, by a 'hellish life' I mean a life at a very low negative welfare level. And by a 'barely bad life' I mean a life at a very high negative welfare level. This allows us to state the

[1] I assume throughout that we are dealing with finite populations that solely contain human persons. See Parfit (1984: §131) for the classic introduction of the Repugnant Conclusion and an argument in its favour. See Ng (1989), Carlson (1998) and Huemer (2008) for other arguments for the Repugnant Conclusion.

Negative Repugnant Conclusion.

For any number of hellish lives, there is a number of barely bad lives that would be worse.

Pace Zuber *et al.* (2021), many continue to believe that we must avoid the (Negative) Repugnant Conclusion. *Non-Archimedean* population axiologies – also called *lexical* views – can do so. For they accept at least one of the following two claims:

The Weak Superiority of Heavenly Lives.

A sufficient number of heavenly lives would be better than any number of barely good lives.[2]

The Weak Inferiority of Hellish Lives.

A sufficient number of hellish lives would be worse than any number of barely bad lives.[3]

The Weak Superiority of Heavenly Lives is inconsistent with the Repugnant Conclusion; and the Weak Inferiority of Hellish Lives is inconsistent with the Negative Repugnant Conclusion. The former captures the intuition that sufficiently large losses in *quality* (welfare) cannot be compensated for by gains in *quantity* (number), no matter how large; and the same goes for the latter, *mutatis mutandis*. Thus, non-Archimedean axiologies are *prima facie* compelling.

Nevertheless, the purpose of this paper is to offer a novel argument against non-Archimedean axiologies. Here's the plan. Section 2 introduces non-Archimedean axiologies in greater detail, showing why they have great initial appeal. Section 3 offers a novel argument against them. To preview, the argument will be that the non-Archimedean solution to the Negative Repugnant Conclusion is untenable, and that without it, the remaining non-Archimedean view entails a different unacceptable claim, which I call the *Repugnant Elitist Conclusion*. Sections 4 and 5 explore the possibilities that are left open if one insists on retaining the non-Archimedean solution to the Negative Repugnant Conclusion. One option, which I call the *Somber View*, turns out to be deeply pessimistic, in that it pushes us towards favouring extinction. The remaining option, which I call *Heavy Tails*, entails a weaker version of the Repugnant Elitist Conclusion. Section 6 considers a final escape route from the argument in section 3. Section 7 concludes.

[2]The Weak Superiority of Heavenly Lives features a relation of *value superiority*. x is *strongly superior* to y iff any quantity of x is better than any quantity of y. x is *weakly superior* to y iff some quantity of x is better than any quantity of y. The Weak Superiority of Heavenly Lives says that heavenly lives are weakly superior to barely good lives. For value superiority proposals, see Parfit (1986, 2016), Griffin (1988: 340 fn. 27), Lemos (1993), Portmore (1999), Kitcher (2000), Rachels (2004), Temkin (2012), Chang (2016), Klocksiem (2016), Thomas (2018), Andersson (2021), Carlson (2022) and Nebel (2022). (Not all of these authors *endorse* these proposals.)

[3]The Weak Inferiority of Hellish Lives features a relation of *value inferiority*. x is *strongly inferior* to y iff any quantity of x is worse than any quantity of y. x is *weakly inferior* to y iff some quantity of x is worse than any quantity of y. The Weak Inferiority of Hellish Lives says that hellish lives are weakly inferior to barely bad lives.

2. The Simple Additive Picture

This section introduces non-Archimedean population axiologies in greater detail. A *population* is a set of lives, each of which has a lifetime welfare level. A *population axiology* is an 'is all-things-considered better than' relation over the set of possible populations. As stated in the Introduction, non-Archimedean axiologies accept at least one of the Weak Superiority of Heavenly Lives and the Weak Inferiority of Hellish Lives. Moreover, as I'll understand them, non-Archimedean axiologies endorse an attractive understanding of the structure of value that I call the *Simple Additive Picture*. The Simple Additive Picture is the conjunction of three principles – *Addition*, *Separability* and *Transitivity* – which I introduce presently.

I begin by building to the principle of Addition. It's plausible that adding a *good life* – a life at a positive welfare level – to a population makes the population better. Why? Huemer (2008: 923) provides a simple rationale: 'Worthwhile lives are good. More of a good thing is better. Therefore, increasing the number of worthwhile lives makes the world better.'[4] Similarly, it's plausible that adding a *bad life* – a life at a negative welfare level – to a population makes the population worse. We can capture these thoughts precisely by introducing some notation. Following Nebel (2023), for all populations A and B, let '$A \cup B$' refer to the union of disjoint populations A^* and B^*, where A^* is identical to A in size and welfare distribution, and likewise for B^* and B. We can now state the principle of

Addition.

For any population X and good (bad) life l, $X \cup l$ is better (worse) than X.

Here's another plausible claim: when we're evaluating a change to a population, people who are entirely unaffected by that change should not factor into our evaluation. Imagine, for example, that we're thinking about adding a new person to Earth's present population. Intuitively, how good or bad it would be to add this person doesn't depend on how many people lived in the Achaemenid Empire (ca. 550–330 BCE), or on how well-off these ancient Persians were. This thought can be captured more precisely in the principle of

Separability.

For all populations X, Y, and Z, X is better than Y iff $X \cup Z$ is better than $Y \cup Z$.[5]

Finally, it's plausible that the 'is all-things-considered better than' relation is transitive. Call this claim *Transitivity*.

[4]Cf. Parfit (2016: 110) on the Simple View, which says that 'Anyone's existence is in itself good, and makes the world in one way better, if this person's life is good to live, or worth living'. See also Broome (2004) for a classic case against the *intuition of neutrality*, the view that it's axiologically neutral to add good lives to the world.

[5]See Thomas (2022b) for an explication and defence of Separability; Goodsell (2021) for an objection to a principle to which Thomas (2022b) appeals in arguing for Separability; and Appendix 1 of this paper for an objection to Variable Value population axiologies, which paradigmatically violate Separability. Cf. Blackorby *et al.*'s (2005) Independence of the Existence of the Dead.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

Addition, Separability and Transitivity comprise the Simple Additive Picture. Now, the most well-known population axiology that endorses the Simple Additive Picture is *Standard Totalism*. According to Standard Totalism, (i) welfare is a scalar quantity (which means that every welfare level can be represented by a single real number) and (ii) for every population A and B, A is better than B just in case total welfare in A is greater than total welfare in B. Standard Totalism infamously entails both the Repugnant Conclusion and the Negative Repugnant Conclusion. Non-Archimedean views consequently enjoy great initial appeal, for, like Standard Totalism, they endorse the Simple Additive Picture, but, unlike Standard Totalism, they are able to avoid the (Negative) Repugnant Conclusion.[6] Moreover, non-Archimedean views can avoid Gustaf Arrhenius's (2011) influential impossibility result in population axiology – at least in cases of choice under certainty.[7] For these reasons, I used to believe that non-Archimedeanism was the best game in town. I have recently come to believe that this was a mistake. The next several sections explain why.

3. The Repugnant Elitist Conclusion

This section lays out my main argument against non-Archimedean axiologies. The argument proceeds in three stages. Firstly, I argue that hellish lives and barely bad lives are *exchangeable*. Loosely, this means that we can trade these types of lives off against each other. (A precise definition is given below.) I argue for this exchangeability claim primarily by arguing against the Weak Inferiority of Hellish Lives. Secondly, I introduce a relation that I call *outweighing* and suggest that non-Archimedeans should accept that barely good lives and barely bad lives can outweigh each other. Finally, given the claims reached thus far, I derive the Repugnant Elitist Conclusion, which I regard as a *reductio*.

3.1. Preliminaries

Two preliminaries before we begin. Firstly, I assume that a population is good (bad) iff it is better (worse) than the *empty population*, which contains zero lives. Secondly, it will be useful to prove up front that the Simple Additive Picture entails a claim that I call *Scaling*. To state Scaling, we'll need to introduce some further notation. For any population X, let '$\mathrm{n}(X)$' refer to a population composed of the union of X and (n-1) *copies* of X, where a copy of a population is a disjoint population identical to the original in size and welfare distribution. Thus, for example, '$2(X)$' refers to $X \cup X^*$, where X^* is identical to X in size and welfare distribution. We can now state

[6]They can do so by understanding welfare as a vector quantity. See Nebel (2022) for discussion.

[7]They can do so by denying an assumption Arrhenius (2011) makes about the structure of welfare called 'Finite Fine-Grainedness'. See Thomas (2018) and Carlson (2022) for discussion of non-Archimedeanism vis-à-vis Arrhenius's (2011) impossibility theorem and Thornley (2021) for a new impossibility theorem for population *prospect* axiology, which non-Archimedeanism cannot avoid.

Scaling.

For all populations A and B, A is better than B iff for every natural number q, $q(A)$ is better than $q(B)$.[8]

We'll start by showing that for all populations A and B, if A is better than B, then for every natural number q, $q(A)$ is better than $q(B)$.[9] Take an arbitrary pair of populations A and B and assume that A is better than B. We want to show that it follows that for every natural number q, $q(A)$ is better than $q(B)$. We'll do so by induction. We already have the base case, where $q = 1$: we've assumed that A is better than B. Here's the inductive step: for every natural number q, if $q(A)$ is better than $q(B)$, then $(q+1)A$ is better than $(q+1)B$. To establish the inductive step, assume that $q(A)$ is better than $q(B)$. It follows from Separability that $q(A) \cup B$ is better than $(q+1)B$. And, given that A is better than B, it also follows from Separability that $(q+1)A$ is better than $q(A) \cup B$. Since $(q+1)A$ is better than $q(A) \cup B$ and $q(A) \cup B$ is better than $(q+1)B$, by Transitivity, $(q+1)A$ is better than $(q+1)B$. This establishes the inductive step, thereby completing the left-to-right half of the proof. Going from right to left is straightforward. Assume that for every natural number q, $q(A)$ is better than $q(B)$. Then in particular A is better than B (this is the special case where $q = 1$).

3.2. Hellish lives and barely bad lives are exchangeable

We now turn to stage one of the argument. The argument will be that hellish lives and barely bad lives are *exchangeable*.

Exchangeability.

One bad (good) is exchangeable with another bad (good) iff for any quantity of the former, there is a quantity of the latter that would be worse (better), and vice versa.

So, if hellish lives and barely bad lives are exchangeable, then for any number of hellish lives, there is a number of barely bad lives that would be worse, and vice versa.

3.2.1. Against the Weak Inferiority of Hellish Lives

To argue that hellish lives and barely bad lives are exchangeable, I first argue against the Weak Inferiority of Hellish Lives. Here's the argument in a nutshell: in conjunction with the Simple Additive Picture, the Weak Inferiority of Hellish Lives entails

[8]Following Thornley (2022), I assume that lives are individuated by the persons leading them. I assume that we are not dealing with fractions of persons; so, every non-empty population contains a natural number of lives. Sometimes, for brevity, I will omit 'natural' in the main text when referring to numbers of lives.

[9]For proof strategy I follow Nebel (2022: 212–213), who proves a different result about heavenly lives and barely good lives.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

Strong Nonsuperiority Across Adjacent Levels.

For any finite, ascending sequence of negative welfare levels beginning with a level that corresponds to a hellish life and ending with a level that corresponds to a barely bad life, there are two welfare levels that are (i) adjacent to one another in the sequence and (ii) such that *no* number of lives at the higher level would be worse than *any* number of lives at the lower level.

Strong Nonsuperiority Across Adjacent Levels is false, so, since we're holding the Simple Additive Picture fixed, we must reject the Weak Inferiority of Hellish Lives.

I'll begin by deriving Strong Nonsuperiority Across Adjacent Levels; then I'll argue against it.[10] It will be useful to have the following two definitions at hand for easy reference:

Weak nonsuperiority.

x is weakly nonsuperior to y iff some quantity of x is not better than any quantity of y.

Strong nonsuperiority.

x is strongly nonsuperior to y iff any quantity of x is not better than any quantity of y.

To save space, I'll refer to a finite, ascending sequence of negative welfare levels beginning with a level that corresponds to a hellish life and ending with a level that corresponds to a barely bad life as a 'negative sequence'. Consider an arbitrary negative sequence S. Call the first level in S 'l_1' and the final level 'l_n'.

We begin by showing that there are two welfare levels in S that are (i) adjacent to one another in S and (ii) such that weak nonsuperiority holds across lives at these levels.[11] Assume for contradiction that there are no such welfare levels. Then for every pair of welfare levels that are adjacent in S, each number of lives at the lower (i.e. worse) welfare level is better than some number of lives at the higher (i.e. better) welfare level. It follows by Transitivity that each number of lives at l_1 is better than some number of lives at l_n. But this contradicts the Weak Inferiority of Hellish Lives, which says that some number of hellish lives (i.e. lives at l_1) is worse than any number of barely bad lives (i.e. lives at l_n). So there are (at least) two welfare levels that are (i) adjacent to one another in S and (ii) such that weak nonsuperiority holds across lives at these levels. Call these welfare levels 'l_j' and 'l_k'.

We'll now show that lives at l_j are strongly nonsuperior to lives at l_k.[12] Assume for contradiction that lives at l_j are not strongly nonsuperior to lives at l_k. This means

[10]Notice that the proof of Strong Nonsuperiority Across Adjacent Levels does *not* assume *Small Steps* (alternatively called 'Finite Fine-Grainedness'), which says that for any two welfare levels, there is a finite sequence of intuitively small differences between them. See Arrhenius (2016), Thomas (2018), Thornley (2022) and Baker (Forthcoming) for discussion of Small Steps.

[11]The sub-proof of this paragraph follows Thornley (2022), who shows the parallel claim in the context of good lives.

[12]The sub-proof of this paragraph is due to Jake Nebel (pers. comm.).

that for some natural numbers m and n, n lives at l_k would be worse than m lives at l_j. Now take an arbitrary natural number x and consider x lives at l_j. There must be a multiple of m – say, qm – such that qm is greater than x. By Addition, qm lives at l_j would be worse than x lives at l_j. However, since n lives at l_k would be worse than m lives at l_j, by Scaling, qn lives at l_k would be worse than qm lives at l_j. Since qn lives at l_k would be worse than qm lives at l_j and qm lives at l_j would be worse than x lives at l_j, by Transitivity, qn lives at l_k would be worse than x lives at l_j. Since x was arbitrary, we have that for any natural number of lives at l_j, there is a natural number of lives at l_k that would be worse. But that contradicts the fact that lives at l_j are weakly nonsuperior to lives at l_k. So lives at l_j are strongly nonsuperior to lives at l_k. And since S was an arbitrary negative sequence, we have Strong Nonsuperiority Across Adjacent Levels.[13]

I'll now give a counterexample to Strong Nonsuperiority Across Adjacent Levels. Here's the setup. We construct a finite sequence of lives that contain nothing but pain. The pain in each life is constant and qualitatively uniform. The lives differ with respect to two variables: duration and intensity of pain. The first life is relatively long and the pain in it is agonizing. It's a hellish life. The final life is one second long and the pain in it is barely noticeable. It's a barely bad life. We can get from the duration of the first life to the duration of the final life via a finite number of small decreases in length, e.g. in increments of 0.5 seconds. Plausibly, we can also get from the intensity of pain in the first life to that in the final life via a finite number of small decreases in pain intensity. We can therefore construct a finite sequence of lives beginning with the first life and ending with the final life such that the sole descriptive difference between lives that are adjacent to one another in the sequence is either a small decrease in duration or a small decrease in pain intensity (but not both). Thus, as I imagine the sequence, the second life contains pain of equal intensity to that in the first life, but is 0.5 seconds shorter; the third life is equal to the second life in duration, but contains pain that is slightly less intense; and so on.

Now consider the sequence of welfare levels that map one-to-one onto the lives in this sequence. This sequence of welfare levels – S' – is a negative sequence. So, according to Strong Nonsuperiority Across Adjacent Levels, there are two welfare levels adjacent to one another in S' such that strong nonsuperiority holds across the lives at these levels. I submit that this is false. For (1) if strong nonsuperiority holds across lives at two different *negative* welfare levels, then the welfare difference between the lives must be very large, in a straightforward intuitive sense. But (2) for every pair of lives that are adjacent in our sequence, it's not the case that the welfare difference between them is very large. So strong nonsuperiority does not hold across

[13]According to *Completeness*, for all populations X and Y, exactly one of the following is true: X is better than Y, Y is better than X, or X and Y are equally good. Given Completeness, strong nonsuperiority collapses into *strong inferiority*, where x is strongly inferior to y iff any quantity of x is worse than any quantity of y. So, if we assume Completeness, Strong Nonsuperiority Across Adjacent Levels will collapse into Strong Inferiority Across Adjacent Levels. Since Strong Inferiority Across Adjacent Levels is even more counterintuitive than Strong Nonsuperiority Across Adjacent Levels, the most plausible non-Archimedean views will allow for incompleteness (so that for some populations A and B, each of the following is false: A is better than B, B is better than A, A and B are equally good). See Nebel (2022) and Thornley (2022) for further discussion.

any pair of lives that are adjacent in our sequence. Therefore, Strong Nonsuperiority Across Adjacent Levels is false.

I'll now defend (1) and (2), in order. Consider an arbitrary pair of negative welfare levels, which we'll call 'level 1' and 'level 2'. If lives at level 1 are strongly nonsuperior to lives at level 2, then no number of lives at level 2 would be worse than any number of lives at level 1. In particular, 10 billion lives at level 2 would not be worse than one life at level 1. For it to be plausible that 10 billion lives at one negative welfare level would not be worse than one life at another negative welfare level, the difference between the welfare levels must be very large, in a straightforward intuitive sense. To illustrate, it is plausible that 10 billion barely bad lives would not be worse than one hellish life. In contrast, it is not plausible that 10 billion terrible-but-not-quite-hellish lives would not be worse than one hellish life. They would be worse.

Turning now to (2): by hypothesis, the descriptive difference between any two lives that are adjacent in our sequence is small. It therefore seems that the welfare difference between any two lives that are adjacent in our sequence is small. However, one might object that small descriptive differences in pain intensity can make a difference to the *type* of pain at issue, which itself is welfare-relevant.[14] For instance, one might hold that although the *descriptive* difference between agony and pain that is terrible-but-not-quite-agony is small, the *welfare* difference between agony and pain that is terrible-but-not-quite-agony is not small. There would then be two lives that are adjacent in our sequence – one containing agony and the other containing pain that is terrible-but-not-quite-agony – such that the welfare difference between them is not small. Is it then plausible to claim that strong nonsuperiority holds across these lives? No. For even if this gambit gives us a rationale for resisting the inference from small descriptive difference to small welfare difference, it does not support the further claim that the welfare difference is *so* large as to ground the holding of strong nonsuperiority across the lives in question. To continue with our example for concreteness: even granting *arguendo* that the welfare difference between the life containing agony and the adjacent life containing pain that is terrible-but-not-quite-agony is not *small*, it does not follow that the welfare difference is *very large* in the sense intended in (1).

Finally, it's worth considering the following partners-in-guilt defence of Strong Nonsuperiority Across Adjacent Levels.[15] Consider a finite, ascending sequence of welfare levels that solely contains some negative welfare levels and a *neutral level* ($=_{\mathrm{df}}$ a welfare level such that adding a life at this level does not make the population better or worse). Suppose that the final two levels in the sequence are a barely negative level and the lone neutral level in the sequence. Many will agree that strong nonsuperiority holds across lives at these two levels, for intuitively, no number of bad lives would be better than any number of neutral lives. What's more, many will retain this judgement even if it is stipulated that the descriptive difference between life at the barely negative level and life at the neutral level is small. By way of illustration, return to the sequence of pain-filled lives that we recently constructed. The final life in this sequence is one

[14]See Klocksiem (2016: 1324) for statement and defence of this position. More generally, one might think that *intrapersonal spectrum arguments* motivate the existence of an evaluatively significant threshold somewhere along the pain continuum; see Pummer (2018) for discussion.

[15]Adapted from Thornley (2022: §4).

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

second long and contains pain that is barely noticeable. Append to this sequence a life that is one second long and hedonically neutral. Plausibly, this is a neutral life; and the descriptive difference between it and its predecessor is small. Thus, the holding of strong nonsuperiority across lives that are descriptively similar is not a unique commitment of non-Archimedeans who endorse the Weak Inferiority of Hellish Lives.

I am not convinced by this partners-in-guilt defence for two reasons. First, the non-Archimedean must claim that strong nonsuperiority holds across different bad lives – e.g. across hellish lives and lives that are terrible-but-not-quite-hellish. In contrast, in claiming that bad lives are strongly nonsuperior to neutral lives (as we just did), our view is that strong nonsuperiority holds across lives that are bad and lives that aren't. This is a particular instance of the more general – and plausible – principle that bad things are strongly nonsuperior to not bad things (i.e. that no amount of a bad thing would be better than any amount of a not bad thing). Second, to avoid the Negative Repugnant Conclusion, non-Archimedeans must claim that there are at least *two* instances of strong nonsuperiority holding across bad lives.[16] As Jensen (2020: 308) observes, this has the unpalatable consequence of creating a 'zone' of incommensurability within the welfare levels, which appears to 'dissolve the difference in value between the members of the zone'.[17] It therefore seems to me false that non-Archimedeans are partners in guilt with those of us who simply wish to claim that bad things are strongly nonsuperior to not bad things.

This concludes my case against Strong Nonsuperiority Across Adjacent Levels.[18] Since the Weak Inferiority of Hellish Lives, in conjunction with the Simple Additive Picture, entails Strong Nonsuperiority Across Adjacent Levels, and we are keeping the Simple Additive Picture fixed, we must reject the Weak Inferiority of Hellish Lives.[19]

[16]Assuming, as I do in footnote 13, that the most plausible non-Archimedean views allow for incompleteness. Rabinowicz (2022: 437 and fn. 12) shows that non-Archimedean views with this structure require at least two instances of strong noninferiority holding across good lives to avoid the Repugnant Conclusion. His argument applies to the Negative Repugnant Conclusion, strong nonsuperiority and bad lives *mutatis mutandis*.

[17]See Jensen (2020: §6) for further discussion.

[18]I have left unaddressed the claim that we can soften the blow of Strong Nonsuperiority Across Adjacent Levels by holding that it is *vague* which welfare levels the strong nonsuperiority relation holds across. I am not optimistic about this move, but lack the space to address it here. See Nebel (2022: §8.4), Pummer (2022) and Thomas (2022c) for discussion.

[19]By deploying arguments parallel to those in this subsection, we can show that in conjunction with the Simple Additive Picture, the Weak Superiority of Heavenly Lives entails *Strong Noninferiority Across Adjacent Levels*: for any finite, descending sequence of positive welfare levels beginning with a level that corresponds to a heavenly life and ending with a level that corresponds to a barely good life, there are two welfare levels that are (i) adjacent to each other in the sequence and (ii) such that no number of lives at the lower level would be better than any number of lives at the higher level. Insofar as we're happy to reject the Weak Inferiority of Hellish Lives on the ground that it entails Strong Nonsuperiority Across Adjacent Levels, why can't we similarly reject the Weak Superiority of Heavenly Lives on the ground that it entails Strong Noninferiority Across Adjacent Levels – and thereby dispense with non-Archimedeanism? I've eschewed this argumentative strategy because I find Strong Nonsuperiority Across Adjacent Levels significantly more counterintuitive than Strong Noninferiority Across Adjacent Levels; and I expect that many readers will share this intuition. Thus, I consider it an advantage of my argument against non-Archimedeanism that it does *not* depend on the intuitive rejection of Strong Noninferiority Across Adjacent Levels.

3.2.2. In favour of exchangeability

If the Weak Inferiority of Hellish Lives is false, does it follow that hellish lives are exchangeable with barely bad lives? No. Handfield and Rabinowicz (2018) and Rabinowicz (2022) prove (assuming Transitivity) that *Trilemma* holds for all goods (bads) that are quantifiable and such that more is better (worse).[20]

Trilemma.

For all such goods (bads) x and y such that x is better (worse) than y, exactly one of the following is true: x is weakly superior (inferior) to y, x and y are exchangeable, or x is *radically incommensurable* with y.

Radical incommensurability.

x is radically incommensurable with y iff there is a quantity k of x such that for every quantity k^+ of x at least as great as k, there is some quantity k' of y such that k^+x is incommensurable with every quantity of y at least as great as k'.

Could hellish lives be radically incommensurable with barely bad lives? If they were, then there would be some natural number k such that for every natural number k^+ at least as great as k, there is a natural number k' such that k^+ hellish lives would be incommensurable with every natural number of barely bad lives at least as great as k'. For concreteness, suppose that $k = 100$ and consider 101 hellish lives. It would follow that there is a natural number – say, 10,000 – such that 101 hellish lives would be incommensurable with every natural number of barely bad lives at least as great as 10,000. Could this be true? I believe the answer is No.

Suppose that hellish lives are radically incommensurable with barely bad lives. Now consider an arbitrary negative sequence and suppose for contradiction that lives at every pair of levels that are adjacent in the sequence are exchangeable. It follows by Transitivity that for any number of lives at the first level (which corresponds to a hellish life), there is a number of lives at the final level (which corresponds to a barely bad life) that would be worse. But that contradicts the supposition that hellish lives are radically incommensurable with barely bad lives. So, it's not the case that lives at every pair of levels that are adjacent in the sequence are exchangeable. It follows by Trilemma that either weak inferiority or radical incommensurability holds across lives at at least one pair of adjacent levels. I assume the former option is a non-starter, so, since the sequence was arbitrary, we have

Radical Incommensurability Across Adjacent Levels.

In any negative sequence, radical incommensurability holds across lives at at least one pair of welfare levels that are adjacent in the sequence.

This is implausible for essentially the same reasons that Strong Nonsuperiority Across Adjacent Levels is implausible. Moreover, the claim that hellish lives are

[20]Strictly speaking, Handfield and Rabinowicz's (2018) proof specifically concerns *harms* and Rabinowicz's (2022) proof specifically concerns *goods*, but their proofs generalize to the Trilemma stated in the main text of this paper.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

radically incommensurable with barely bad lives does not even enjoy the intuitive appeal of the Weak Inferiority of Hellish Lives. I conclude against the radical incommensurability of hellish and barely bad lives. By Trilemma, we are left with exchangeability.

3.3. Outweighing

We now proceed to stage two of the argument. Here, I introduce the *outweighing* relation and suggest that non-Archimedeans should accept that barely good lives and barely bad lives can outweigh each other.

> **Outweighing.**
>
> x can outweigh y iff for any quantity k of y that is good (bad), there is a quantity k' of x such that ky together with k'x is bad (good).

Non-Archimedeans should accept that barely bad lives can outweigh barely good lives, because denying this claim is intuitively intolerable. Here's why. Suppose that barely bad lives could not outweigh barely good lives. Then there would be some natural number n such that n barely good lives along with *any* number of barely bad lives would not be a bad population. That is unacceptable. n barely good lives along with 10^n barely bad lives *would* be a bad population. So barely bad lives can outweigh barely good lives.

Can barely good lives outweigh barely bad lives? If they could not, then there would be some natural number m such that m barely bad lives along with any number of barely good lives would not be a good population. To me, this does not register as obviously false in the way that the corresponding claim about good lives does. However, we can show a stronger result: if m barely bad lives along with any number of barely good lives would not be a good population, then a *single* barely bad life along with any number of barely good lives would not be a good population. To see how, assume for contradiction that there is some natural number x such that x barely good lives ∪ one barely bad life would be a good population, i.e. would be better than the empty population. It follows from Scaling that for every natural number q, qx barely good lives ∪ q barely bad lives would be better than the empty population. But that is inconsistent with the claim that barely good lives can't outweigh barely bad lives – i.e. that there is a natural number m such that m barely bad lives ∪ any number of barely good lives would *not* be good, i.e. would not be better than the empty population. So, if barely good lives can't outweigh barely bad lives, then a single barely bad life cannot be taken together with any number of barely good lives to form a good population.

However, it seems *internally* ill-motivated for non-Archimedeans to deny that there is some number of barely good lives that can be taken together with a single barely bad life to form a good population. Despite being barely so, barely good lives are good. They are positively valuable for the persons leading them. By Addition, the addition of each barely good life makes the world better. In virtue of what, then, would the goodness of enough such lives – we can make the number as large as we like – fail to overcome the badness of a single barely bad life? For now, I'll assume

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

that non-Archimedeans lack a compelling answer to this question, but we'll revisit this assumption in section 6.[21]

3.4. The Repugnant Elitist Conclusion

So far, I've argued that hellish lives and barely bad lives are exchangeable (stage 1) and that non-Archimedeans should accept that barely good lives and barely bad lives can outweigh one another (stage 2). We can now derive the

> **Repugnant Elitist Conclusion.**
>
> Some number of heavenly lives along with any number of hellish lives would be a good population.

We begin by showing that for all good populations X and Y and every bad population Z, if X is better than Y and $Y \cup Z$ is good, then $X \cup Z$ is good too. To do so, assume that we have two arbitrary good populations X and Y such that X is better than Y and an arbitrary bad population Z such that $Y \cup Z$ is good, i.e. better than the empty population. Since X is better than Y, by Separability, $X \cup Z$ is better than $Y \cup Z$. Since $X \cup Z$ is better than $Y \cup Z$ and $Y \cup Z$ is better than the empty population, by Transitivity, $X \cup Z$ is better than the empty population – i.e. $X \cup Z$ is good. Using the same strategy, we can also show that for all bad populations X and Y and every good population Z, if X is worse than Y and $X \cup Z$ is good, then $Y \cup Z$ is good too. The proof is omitted for brevity.

Now for the Repugnant Elitist Conclusion:

1. There is a natural number t such that t heavenly lives would be better than any natural number of barely good lives. (Weak Superiority of Heavenly Lives)
2. For every natural number of barely bad lives, there is a natural number of barely good lives such that the barely bad lives $\cup$ the barely good lives would be a good population. (Barely good lives can outweigh barely bad lives)
3. For all good populations X and Y and every bad population Z, if X is better than Y and $Y \cup Z$ is good, then $X \cup Z$ is good.

[21]On the standard conception of the welfare levels, the welfare levels consist exhaustively of positive welfare levels, negative welfare levels and a single neutral level. In contrast, Gustafsson (2020) explores an alternative picture on which there are positive welfare levels, negative welfare levels and multiple *undistinguished* welfare levels, where life at an undistinguished welfare level is neither good, nor bad, nor neutral for the person leading it. An anonymous referee suggests that Gustafsson's alternative picture undermines intuitive support for the claim that barely good lives can outweigh barely bad lives, because, on the alternative picture, barely bad lives may be significantly worse than barely good lives. However, even granting Gustafsson's alternative picture *arguendo*, it's not clear to me why the fact that barely bad lives are significantly worse than barely good lives would – in itself – undermine the outweighing thesis of §3.3. After all, if barely bad lives are significantly worse than barely good lives, then barely good lives are significantly better than barely bad lives. What *would* undermine the outweighing thesis is the claim that barely bad lives are (significantly) more bad than barely good lives are good – i.e. that barely bad lives have a (significantly) greater absolute value than barely good lives. But as far as I am aware, this is a substantive further claim that (i) does not simply fall out of Gustafsson's alternative picture and (ii) is unmotivated, at least as things stand at present.

4. t heavenly lives ∪ any number of barely bad lives would be a good population. (1–3)
5. For any natural number of hellish lives, there is a natural number of barely bad lives that would be worse. (Exchangeability)
6. For all bad populations X and Y and every good population Z, if X is worse than Y and $X \cup Z$ would be good, then $Y \cup Z$ would be good.
7. t heavenly lives ∪ any number of hellish lives would be a good population. (4–6)

(7) is the Repugnant Elitist Conclusion, which I regard as a *reductio*. (Intuitively, for any natural number of heavenly lives, there is some natural number of hellish lives such that the heavenly and hellish lives together would not be a good population.) To drive a final nail into the coffin, we can also show that one heavenly life ∪ any natural number of hellish lives would not be a bad population. To see how, assume for contradiction that there is a natural number n such that one heavenly life ∪ n hellish lives would be bad, i.e. worse than the empty population. Then, by Scaling, for every natural number q, q heavenly lives ∪ qn hellish lives would be worse than the empty population. But that contradicts the Repugnant Elitist Conclusion.

3.5. Argument recap

Non-Archimedean population axiologies endorse the Simple Additive Picture and the Weak Superiority of Heavenly Lives and/or the Weak Inferiority of Hellish Lives. We can name and taxonomize the resulting positions as follows in Table 1.

Table 1. Non-Archimedean population axiologies

	Weak Superiority of Heavenly Lives	Weak Inferiority of Hellish Lives
The Rosy View	✓	✗
The Somber View	✗	✓
Heavy Tails	✓	✓

Here's the main argument against non-Archimedeanism in a nutshell: in conjunction with the Simple Additive Picture, the Weak Inferiority of Hellish Lives entails Strong Nonsuperiority Across Adjacent Levels. That's false, so we should reject the Weak Inferiority of Hellish Lives. That rules out the Somber View and Heavy Tails but leaves open the Rosy View. However, the Rosy View entails the Repugnant Elitist Conclusion, which is false. So all non-Archimedean positions fail.

The next few sections proceed as follows: in due recognition of the fact that not everyone will find Strong Nonsuperiority Across Adjacent Levels as damning as I do, I consider the Somber View in the next section and Heavy Tails in the one after. Then, in section 6, I consider the option of denying that barely good lives can outweigh barely bad lives.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

4. The Somber View

The Somber View is the mirror image of the Rosy View. It accepts the Weak Inferiority of Hellish Lives and the claim that barely good lives and barely bad lives can outweigh each another. And, just as the Rosy View holds that hellish lives and barely bad lives are exchangeable, the Somber View holds that heavenly lives and barely good lives are exchangeable. Thus, just as the Rosy View bites the bullet on the Negative Repugnant Conclusion, the Somber View bites the bullet on the Repugnant Conclusion.

We can run an argument parallel to the one given for the Repugnant Elitist Conclusion to show that the Somber View implies the following: any population with a sufficient number of hellish lives is bad, no matter the number of heavenly lives it also contains (see Appendix 2 for proof).[22] I find this implication *prima facie* plausible, unlike the Repugnant Elitist Conclusion. For it is plausible that the extremes of suffering are so horrific that an outcome is irredeemably marred if it contains enough of them.[23] That is the chief reason why the Somber View is worthy of note and discussion. However, in light of the grim inductive evidence from the historical record and the many different ways in which the future could play out, there is, disturbingly, a large expected number of hellish lives in the future.[24] Now, strictly speaking, the Somber View does not say anything about the future, for, as stated, it does not say anything about the evaluation of risky prospects. But a natural synthesis of the Somber View with expected value reasoning will yield the result that the future is bad in expectation.[25] So, on the Somber View, it would be expectedly

[22]Mogensen (Forthcoming) discusses *Lexical Threshold Negative Utilitarianism* (LTNU), which is stronger than the Somber View. According to LTNU, *one* hellish life along with any number of good lives would be a bad population. Mogensen derives this conclusion from Separability, Transitivity, the Weak Inferiority of Hellish Lives (which he calls the 'Contrary Reverse Repugnant Conclusion'), and two claims to which I do not appeal: first, Completeness, which says that for all populations *A* and *B*, exactly one of the following is true: *A* is better than *B*, *B* is better than *A*, or *A* and *B* are equally good; and second, that bad lives can outweigh good lives. (In contrast, I claim – more weakly – that barely bad lives can outweigh barely good lives). Mogensen later shows that if one drops Completeness, one can derive *Weak Lexical Threshold Negative Utilitarianism* (WLTNU), which, like the Somber View, implies that once a population has a sufficient number of hellish lives, it's bad. He does not consider the objection that (W)LTNU implies Strong Nonsuperiority Across Adjacent Levels, however.

[23]Crisp (2021) is sympathetic to this perspective.

[24]As Edward Gibbon is famously quoted: 'history … is, indeed, little more than the register of the crimes, follies, and misfortunes of mankind' (Gibbon 1996 [1776]: Vol. 1, Ch. 3, pt. 2). Gibbon's quip is undoubtedly hyperbolic, but that it's humorous, rather than simply puzzling, testifies to the element of truth it contains. See, e.g. Glover (2012) on the horrors contained in just one century of human history (the 20th) and Caviola *et al.* (2021, reproduced in MacAskill 2022: 200), who find, in a survey administered to 240 people in India and 240 people in the United States, that 13.5% of respondents believe that their lives have contained more suffering than happiness, 7.7% would prefer to have never been born, and 25% would not relive their lives. As far as possible futures go, all of the following – in which hellish lives would very likely be lived – are live epistemic possibilities: dystopia (e.g. a long-lasting global totalitarianism), the unrecoverable collapse of industrial civilization (perhaps due to a nuclear WWIII), and long-term economic stagnation. See Ord (2020) and MacAskill (2022) for discussion of possible future scenarios.

[25]The Somber View implies that any population with at least some number x of hellish lives is bad. So, on the synthesis I'm imagining, if the expected number of hellish lives in a future population is at least as great as x, then that future population is bad in expectation. Given that there is a large expected number of hellish lives in the future, the future is expectedly bad by the lights of the Somber View – at least in terms of human

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

better if we succumbed to an immediate, painless extinction. That is a difficult pill to swallow.[26] Moreover, the Somber View implies that no number of heavenly lives along with a single hellish life would be a good population (see Appendix 2 for proof). The Somber View therefore has two options: either a single hellish life along with any number of heavenly lives would constitute a bad population, or it would constitute a population that is neither good, nor bad, nor neutral – i.e. a population that is incommensurable with the empty population. Neither option seems right. For a population containing one hellish life and 10 trillion heavenly lives seems good overall.[27]

5. Heavy Tails

The preceding observations suggest to me that the Somber View is false. It remains for us to examine Heavy Tails. Heavy Tails affirms the Weak Superiority of Heavenly Lives; the Weak Inferiority of Hellish Lives; that barely good lives and barely bad lives can outweigh each other; and that heavenly and hellish lives can outweigh each other.[28] Heavy Tails thus avoids both the Repugnant Elitist Conclusion and the Somber View's proclivities towards extinction – in addition to both the Repugnant Conclusion and the Negative Repugnant Conclusion. On this basis, Heavy Tails seems to me the *relatively* most plausible non-Archimedean view, despite the fact that it bites the bullet on Strong Nonsuperiority Across Adjacent Levels. However, Heavy Tails implies that one heavenly life along with any number of barely bad lives would not be a bad population (see Appendix 2 for proof). This implication – call it the *Weak Repugnant Elitist Conclusion* – is rather counterintuitive. For it seems that a population containing one heavenly life and 10 trillion barely bad lives *would* be bad (for instance). Heavy Tails must deny this. To be thorough, though, it is also worth noting that Heavy Tails has the parallel implication that one hellish life along with any number of barely good lives would not be a good population. I lack a strong intuition about this implication, though I appreciate that some will find it welcome.

welfare, and barring the ad hoc manoeuvre of stipulating that x is so high as to be practically unreachable (a move that Mogensen (Forthcoming) also dismisses; cf. Mogensen as well on setting the lifetime welfare level of a hellish life so low as to be practically unrealizable).

[26]Similarly, Mogensen (Forthcoming) writes, 'the greatest cost of accepting [Lexical Threshold Negative Utilitarianism, on which see footnote 22] is surely that it appears to support the desirability of human extinction'.

[27]For agreement, see Mogensen's (Forthcoming) assessment of Ursula K. Le Guin's 'The Ones Who Walk Away from Omelas'.

[28]Why assume that heavenly lives can outweigh hellish lives? Two reasons. First, it gives Heavy Tails a natural, appealing symmetry. Second, assume that they can't. Then there's a natural number n such that n hellish lives along with any number of heavenly lives wouldn't be good. Since there is in expectation a large number of hellish lives in the future (see §4 above), we expect that the future will not be good, no matter how many heavenly lives it contains (at least in welfarist terms, and again, barring the ad hoc manoeuvre of setting n so high that it could never be reached in practice – see footnote 25). I am unprepared to accept this implication, though I grant that this is a reasonable point of disagreement.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

6. Outweighing Redux

Might the preceding arguments be taken as a *reductio* of the claim that barely good lives and barely bad lives can outweigh each other (§3.3)? I'll take the claim that barely bad lives can outweigh barely good lives as unassailable, since it's extremely plausible that one barely good life can be taken together with *some* number of barely bad lives to form a bad population. In contrast, the claim that one barely bad life can be taken together with some number of barely good lives to form a good population enjoys less intuitive support. So, the place to push is against the claim that barely good lives can outweigh barely bad lives.

The non-Archimedean could alternatively claim that one barely bad life along with any number of barely good lives would be neither good, nor bad, nor neutral (i.e. that it would be neither better than, nor worse than, nor equally as good as the empty population, but rather incommensurable with the empty population). This move continues to strike me as ad hoc, given that, by the non-Archimedean's own lights, the world just gets better and better as we add (barely) good lives.[29] Still, it does have the welcome effect of blocking the arguments for the Repugnant Elitist Conclusion and the Weak Repugnant Elitist Conclusion; and for this reason, I expect that many non-Archimedeans will go in for it. What, then, is the best view for the non-Archimedean who takes this route?

Denying that barely good lives can outweigh barely bad lives does not help the Somber View, so we'll examine revisions of Heavy Tails and the Rosy View. *Revised Heavy Tails* is the same as Heavy Tails, except that it denies that barely good lives can outweigh barely bad lives – thereby avoiding the Weak Repugnant Elitist Conclusion. Unfortunately, Revised Heavy Tails – alongside Heavy Tails – implies that there is a natural number n such that n heavenly lives $\cup$ any natural number of barely bad lives would be a good population (see Appendix 2 for proof).[30] This claim – call it *Heavenly Dominance* – is counterintuitive. For instance, it seems that for every natural number q, q heavenly lives $\cup$ 10^{100q} barely bad lives would *not* be a good population. As Mogensen (Forthcoming) remarks, 'it may well strike us most repugnant of all to assert that there are some lives so good that, for their sake, we should be willing to accept that arbitrarily many individuals may have to have lives bad enough that, for their sake, we should wish that they had never been born'. It's also worth noting, though, that (Revised) Heavy Tails has the parallel implication that there is a natural number m such that m hellish lives $\cup$ any natural number of barely good lives would be a bad population. I lack a clear intuition about this implication, but I expect that many will find it welcome.

A final option to consider is the *Revised Rosy View*. The Revised Rosy View is the same as the Rosy View, except that it denies that barely good lives can outweigh barely bad lives – thereby avoiding the Repugnant Elitist Conclusion. Moreover,

[29]Though see Hájek and Rabinowicz (2022) for a possible rationale.

[30]Cf. Mogensen (Forthcoming) on *Lexical Total Utilitarianism*, which is very similar, if not identical, to Revised Heavy Tails and which also entails Heavenly Dominance. I've assumed throughout that the populations we consider solely contain human lives. However, Mogensen has an interesting proposal – which he does not endorse – for why it might be desirable for a more general population axiology to accept Heavenly Dominance: we might feel that a world containing a sufficient number of heavenly *human* lives would be good, even if it contained arbitrarily many more barely bad *non-human animal* lives – e.g. lives of insects and fish that are just barely negative in welfare. I lack this intuition and worry that whatever intuitive appeal the example may have rests on an objectionable form of speciesism.

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

unlike Revised Heavy Tails, the Revised Rosy View does not imply Strong Nonsuperiority Across Adjacent Levels or Heavenly Dominance. It is the ability of the Revised Rosy View to avoid these two unattractive implications of Revised Heavy Tails that makes it a live option. (Otherwise, one might wonder about the motivation for a non-Archimedean view that avoids the Repugnant Conclusion but fails to avoid the Negative Repugnant Conclusion.)

7. Conclusion

The main goal of this paper has been to offer a novel argument against non-Archimedean population axiologies, which I did in section 3. A secondary goal has been to assess the relative plausibility of competing non-Archimedean axiologies, which I attempted to do by identifying their most counterintuitive implications. Results are summarized in Tables 2 and 3.

Table 2. Non-Archimedean population axiologies (expanded)

	Weak Superiority of Heavenly Lives	Weak Inferiority of Hellish Lives	Barely good lives can out-weigh barely bad lives
Rosy View	✓	✗	✓
Revised Rosy View	✓	✗	✗
Somber View	✗	✓	✓
Heavy Tails	✓	✓	✓
Revised Heavy Tails	✓	✓	✗

Table 3. Itemized billing for non-Archimedean axiologies

View	Key costs
Rosy View	1. Negative Repugnant Conclusion 2. Repugnant Elitist Conclusion
Revised Rosy View	1. Negative Repugnant Conclusion 2. One barely bad life along with any number of barely good lives would not be a good population (problem of internal motivation)
Somber View	1. Repugnant Conclusion 2. Strong Nonsuperiority Across Adjacent Levels 3. No number of heavenly lives along with one hellish life would be a good population 4. Tends to favour extinction – including in the actual world
Heavy Tails	1. Strong Nonsuperiority Across Adjacent Levels 2. Weak Repugnant Elitist Conclusion 3. Heavenly Dominance
Revised Heavy Tails	1. Strong Nonsuperiority Across Adjacent Levels 2. One barely bad life along with any number of barely good lives would not be a good population (problem of internal motivation) 3. Heavenly Dominance

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

To close, I'd like to make three conjectures that I hope will be productively provocative. Firstly, if there is such a thing as *the correct abductive methodology for ethical theorizing*, and if this methodology gives significant weight to simplicity – as abduction seems to in scientific theorizing – then we ought to prefer Standard Totalism to non-Archimedeanism. Standard Totalism is much simpler than non-Archimedeanism. Still, it has seemed to many that non-Archimedeanism is more extensionally adequate than Standard Totalism,[31] with the chief worry levelled against the former being its implications for choice under risk.[32] I take the results shown in this paper to narrow the gap in extensional adequacy between non-Archimedeanism and Standard Totalism – even in cases of choice under certainty. With a sufficiently narrow gap in extensional adequacy, considerations of simplicity can then carry the all-things-considered judgement of which view to prefer.

Secondly, if Standard Totalism is true, then familiar forms of maximizing welfarist consequentialism are false. According to Standard Totalism, all good lives are exchangeable, all bad lives are exchangeable, and good and bad lives can all outweigh each other. Standard Totalism implies that for any number of hellish lives, there is a number of barely good lives such that it would be better to add the barely good lives to the world than to alleviate the suffering contained in the hellish lives (Temkin 2012: 413). Although I am prepared to entertain that this is true as a matter of impersonal axiology – i.e. as a matter of what would be better 'from the point of view of the universe' – I am unprepared to accept that we morally ought to bring into existence some large number of people with barely good lives, rather than alleviating the extreme suffering of people who already exist.

Disturbingly, however, we can pose the paradoxical arguments in population *axiology* that lead us to Standard Totalist repugnance in *deontic* terms (Arrhenius 2004). These deontic arguments constitute a major challenge to nonconsequentialist moral theories. As I see things, then, one of the most important open questions in ethical theory is whether nonconsequentialists can meet this challenge. If they could not, then I would be inclined towards the view that the impossibility of population ethics supports either moral scepticism or metaethical error theory.[33] The error theory would, of course, undermine this essay, for it would imply that population ethics has no subject matter. Nevertheless, *reasoning through* the paradoxes of population ethics would still have been worthwhile, insofar as it was instrumental in allowing us to grasp the metaethical truth or to better understand our own (mind-dependent) values.

[31]In addition to the Repugnant Conclusion and the Negative Repugnant Conclusion, Standard Totalism implies the *Very Repugnant Conclusion*: for any number of heavenly lives, and for any *greater* number of hellish lives, the hellish lives along with some number of barely good lives would be better than the heavenly lives (Arrhenius 2011: 2).

[32]See Nebel (2022: §8.5) and Thomas (2022a: 495) for worries about non-Archimedeanism under risk. Nebel (2022: 221) calls this 'the most serious problem for [non-Archimedean] views'.

[33]See Cowie (Unpublished manuscript) and Cowie (2022) for just such arguments in favour of scepticism and error theory, respectively. See also McMahan (2013: 34), who writes that 'It is these problems [*viz*., the problems of population ethics] … rather than arguments in metaethics about the queerness of objective values, the connections between normativity and motivation, and so on, that seem to me to pose the greatest challenge to realism in ethics'.

Acknowledgements. I am grateful to Lara Buchak, Pietro Cibinel, Adam Elga, Sam Fullhart, Harvey Lederman, Gideon Rosen, Asher Shang and two anonymous referees for valuable written comments on earlier drafts of this article and to Sebastian Liu, an audience at Princeton University, and an audience at Rocky Mountain Ethics Congress XVI for valuable discussion. My greatest debt is to Jake Nebel, who offered detailed feedback on multiple drafts of the article.

Funding statement. None.

Competing interests. None.

References

Andersson H. 2021. Spectrum arguments, indeterminacy, and value superiority. In *Value Incommensurability: Ethics, Risk, and Decision-Making*, ed. H. Andersson and A. Herlitz, 109–125. New York, NY: Routledge.

Arrhenius G. 2004. The paradoxes of future generations and normative theory. In *The Repugnant Conclusion: Essays on Population Ethics*, ed. T. Tännsjö and J. Ryberg, 201–218. Dordrecht: Kluwer Academic.

Arrhenius G. 2016. Population ethics and different-number-based imprecision. *Theoria* **82**, 166–181.

Arrhenius G. 2011. The impossibility of a satisfactory population ethics. In *Descriptive and Normative Approaches to Human Behavior. Advanced Series on Mathematical Psychology*, ed. E. Dzhafarov and L. Perry, 1–26. Singapore: World Scientific Publishing Company.

Asheim G.B. and S. Zuber 2014. Escaping the repugnant conclusion: rank discounted utilitarianism with variable population. *Theoretical Economics* **9**, 629–650.

Baker C. Forthcoming. Theravāda Buddhism, Small Steps and the Repugnant Conclusion. *Journal of Buddhist Ethics*.

Blackorby C., W. Bossert and D. Donaldson 2005. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press.

Broome J. 2004. *Weighing Lives*. Oxford: Oxford University Press.

Carlson E. 1998. Mere addition and two trilemmas of population ethics. *Economics & Philosophy* **14**, 283–306.

Carlson E. 2022. On some impossibility theorems in population ethics. In *The Oxford Handbook of Population Ethics*, ed. G. Arrhenius, K. Bykvist, T. Campbell and E. Finneron-Burns, 204–225. New York, NY: Oxford University Press.

Caviola L., A. Hoskin and J. Lewis 2021. Do people think their lives are worth living? A survey in India and the US. Unpublished MS.

Chang R. 2016. Parity, imprecise comparability, and the repugnant conclusion. *Theoria* **82**, 183–215.

Cowie C. 2022. A new argument for moral error theory. *Noûs* **56**, 276–294.

Cowie C. The unsolvable puzzle: a new case for moral scepticism. Unpublished manuscript.

Crisp R. 2021. Would extinction be so bad? *New Statesman*. https://www.newstatesman.com/ideas/agora/2021/08/would-extinction-be-so-bad.

Gibbon E. 1996 [1776]. *The History of the Decline and Fall of the Roman Empire*. Project Gutenberg. https://www.gutenberg.org/files/25717/25717-h/25717-h.htm.

Glover J. 2012. *Humanity: A Moral History of the Twentieth Century*. New Haven, CT: Yale University Press.

Goodsell Z. 2021. A St Petersburg paradox for risky welfare aggregation. *Analysis* **81**, 420–426.

Griffin J. 1988. *Well-Being: Its Meaning, Measurement, and Moral Importance*. Oxford: Oxford University Press.

Gustafsson J.E. 2020. Population axiology and the possibility of a fourth category of absolute value. *Economics & Philosophy* **36**, 81–110.

Hájek A. and W. Rabinowicz 2022. Degrees of commensurability and the repugnant conclusion. *Noûs* **56**, 897–919.

Handfield T. and W. Rabinowicz 2018. Incommensurability and vagueness in spectrum arguments: options for saving the transitivity of betterness. *Philosophical Studies* **175**, 2373–2387.

Huemer M. 2008. In defence of repugnance. *Mind* **117**, 899–933.

Hurka T. 1983. Value and population size. *Ethics* **93**, 496–507.
Jensen K.K. 2020. Weak superiority, imprecise equality and the repugnant conclusion. *Utilitas* **32**, 294–315.
Kitcher P. 2000. Parfit's puzzle. *Noûs* **34**, 550–577.
Klocksiem J. 2016. How to accept the transitivity of *better than*. *Philosophical Studies* **173**, 1309–1334.
Lemos N.M. 1993. Higher goods and the myth of Tithonus. *Journal of Philosophy* **90**, 482–96.
MacAskill W. 2022. *What We Owe the Future*. New York, NY: Basic Books.
McMahan J. 2013. Causing people to exist and saving people's lives. *Journal of Ethics* **17**, 5–35.
Mogensen A. Forthcoming. The weight of suffering. *Journal of Philosophy*. GPI Working Paper No. 4-2022.
Nebel J.M. 2022. Totalism without repugnance. In *Ethics and Existence: The Legacy of Derek Parfit*, ed. J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, 200–231. Oxford: Oxford University Press.
Nebel J.M. 2023. The sum of well-being. *Mind*. https://doi.org/10.1093/mind/fzad019.
Ng Y.K. 1989. What should we do about future generations? Impossibility of Parfit's theory X. *Economics & Philosophy* **5**, 235–253.
Ord T. 2020. *The Precipice: Existential Risk and the Future of Humanity*. London: Bloomsbury.
Parfit D. 1984. *Reasons and Persons*. Oxford: Oxford University Press.
Parfit D. 1986. Overpopulation and the quality of life. In *Applied Ethics*, ed. P. Singer, 145–164. Oxford: Oxford University Press.
Parfit D. 2016. Can we avoid the repugnant conclusion?" *Theoria* **82**, 110–127.
Pivato M. 2020. Rank-additive population ethics. *Economic Theory* **69**, 861–918.
Portmore D.W. 1999. Does the total principle have any repugnant implications?" *Ratio* **12**, 80–98.
Pummer T. 2018. Spectrum arguments and hypersensitivity. *Philosophical Studies* **175**, 1729–1744.
Pummer T. 2022. Sorites on what matters. In *Ethics and Existence: The Legacy of Derek Parfit*, ed. J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, 498–523. Oxford: Oxford University Press.
Rabinowicz W. 2022. Can Parfit's appeal to incommensurabilities in value block the continuum argument for the repugnant conclusion?" In *Ethics and Existence: The Legacy of Derek Parfit*, ed. J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, 430–460. Oxford: Oxford University Press.
Rachels S. 2004. Repugnance or intransitivity: a repugnant but forced choice. In *The Repugnant Conclusion: Essays on Population Ethics*, ed. J. Ryberg and T. Tännsjö, 163–186. Dordrecht: Kluwer Academic.
Sider T.R. 1991. Might theory X be a theory of diminishing marginal value? *Analysis* **51**, 265–271.
Temkin L.S. 2012. *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*. New York, NY: Oxford University Press.
Thomas T. 2018. Some possibilities in population axiology. *Mind* **127**, 807–832.
Thomas T. 2022a. On evaluative imprecision. In *Ethics and Existence: The Legacy of Derek Parfit*, ed. J. McMahan, T. Campbell, J. Goodrich and K. Ramakrishnan, 478–497. Oxford: Oxford University Press.
Thomas T. 2022b. Separability and population ethics. In *the Oxford Handbook of Population Ethics*, ed. G. Arrhenius, K. Bykvist, T. Campbell and E. Finneron-Burns, 271–295. New York, NY: Oxford University Press.
Thomas T. 2022c. Are spectrum arguments defused by vagueness? *Australasian Journal of Philosophy* **100**, 743–757.
Thornley E. 2021. The impossibility of a satisfactory population prospect axiology (independently of Finite Fine-Grainedness). *Philosophical Studies* **178**, 3671–3695.
Thornley E. 2022. A dilemma for lexical and Archimedean views in population axiology. *Economics & Philosophy* **38**, 395–415.
Zuber S. *et al.* 2021. What should we agree on about the repugnant conclusion? *Utilitas* **33**, 379–383.

Appendix 1. Separability and Variable Value Theories

To my mind, Separability is the least plausible of the three principles that comprise the Simple Additive Picture.

Separability.
For all populations X, Y, and Z, X is better than Y iff $X \cup Z$ is better than $Y \cup Z$.

Unfortunately, it has proven difficult to formulate a plausible population axiology that violates Separability. Perhaps the most well-known axiologies that do so are *Variable Value theories* (on which see Hurka 1983;

Ng 1989; Sider 1991; Asheim and Zuber 2014; Pivato 2020). According to Variable Value theories, good lives have diminishing marginal value. However, it's extremely plausible that bad lives have non-diminishing marginal disvalue; and difficulties arise if we believe that there's an asymmetry in the marginal (dis)value of good and bad lives. Firstly, we incur an explanatory burden: what explains the asymmetry? Secondly, we're saddled with the following strongly counterintuitive implication. Imagine that God creates a population consisting of an arbitrarily large number of heavenly lives and one barely bad life. Intuitively, this population is good (i.e. better than the empty population). However, if God iteratively creates new populations with identical welfare distributions (we can imagine that this happens on causally isolated planets), then the universe will eventually become bad and then get ever-increasingly so. This is because the value of the heavenly lives will approach a finite upper bound as their number approaches infinity, while the disvalue of the barely bad lives will grow without limit. Eventually, the unbounded disvalue must eclipse the bounded value. So, by successively adding intrinsically good, causally isolated populations to the universe, we can make the universe an arbitrarily bad place. I find this mysterious.[34] It also has dire implications for the future of humanity. For unless we are sanguine about our chances of building a utopia in which there are few bad lives, the long-term survival of humanity is likely to be bad (in expectation, and in welfarist terms) on account of the asymmetry between good and bad lives.

Appendix 2. Proofs

Proof: the Somber View implies that some number of hellish lives along with any number of heavenly lives would be a bad population.

1. There is a natural number t such that t hellish lives would be worse than any natural number of barely bad lives. (Weak Inferiority of Hellish Lives)
2. For any natural number of barely good lives, there is a natural number of barely bad lives such that the barely good lives $\cup$ the barely bad lives would be a bad population. (Barely bad lives can outweigh barely good lives)
3. For all bad populations X and Y and every good population Z, if X is worse than Y and $Y \cup Z$ would be bad, then $X \cup Z$ would be bad.
4. t hellish lives $\cup$ any natural number of barely good lives would be a bad population. (1–3)
5. For any natural number of heavenly lives, there is a natural number of barely good lives that would be better. (Exchangeability of good lives)
6. For all good populations X and Y and every bad population Z, if X is better than Y and $X \cup Z$ would be bad, then $Y \cup Z$ would be bad.
7. t hellish lives $\cup$ any natural number of heavenly lives would be a bad population. (4–6)

Proof: the Somber View implies that no natural number of heavenly lives along with a single hellish life would be a good population.

Assume for contradiction that there is a natural number n such that n heavenly lives $\cup$ one hellish life would be a good population. By Scaling, for every natural number q, qn heavenly lives $\cup$ q hellish lives would be a good population. But this contradicts the fact that on the Somber View, there is a natural number t such that t hellish lives $\cup$ any natural number of heavenly lives would be a bad population.

Proof: Heavy Tails implies that one heavenly life $\cup$ any number of barely bad lives would not be a bad population (i.e. the Weak Repugnant Elitist Conclusion).

We begin by showing that there is a natural number t such that t heavenly lives $\cup$ any number of barely bad lives would be a good population. This is simply the first half of the proof of the Repugnant Elitist Conclusion:

[34]This objection is closely related to, but distinct from, Parfit's Ridiculous and Absurd Conclusions (1984: §138). (The Ridiculous Conclusion takes the unit of aggregation to be discrete experiences, whereas I aggregate over lives (as throughout). The Absurd Conclusion trades on difficulties that arise from capping the quantity of value that can be realized within a given period of time, whereas I adopt a 'timeless' perspective.)

https://doi.org/10.1017/S0266267124000099 Published online by Cambridge University Press

1. There is a natural number t such that t heavenly lives would be better than any natural number of barely good lives. (Weak Superiority of Heavenly Lives)
2. For any natural number of barely bad lives, there is a natural number of barely good lives such that the barely bad lives ∪ the barely good lives would be a good population. (Barely good lives can outweigh barely bad lives)
3. For all good populations X and Y and every bad population Z, if X is better than Y and $Y \cup Z$ is good, then $X \cup Z$ is good.
4. t heavenly lives ∪ any number of barely bad lives would be a good population. (1–3)

Now assume for contradiction that there is a natural number n such that n barely bad lives ∪ one heavenly life would be a bad population. Then, by Scaling, for every natural number q, qn barely bad lives ∪ q heavenly lives would be a bad population. But this contradicts the fact that there is a natural number t such that t heavenly lives ∪ any natural number of barely bad lives would be a good population. So there is no natural number n such that n barely bad lives ∪ one heavenly life would be a bad population.

Proof: (Revised) Heavy Tails implies that some number of heavenly lives ∪ any number of barely bad lives would be a good population (i.e. Heavenly Dominance).

The Weak Inferiority of Hellish Lives means that there is a natural number m such that m hellish lives would be worse than any natural number of barely bad lives. And since heavenly lives can outweigh hellish lives, there is a natural number n such that n heavenly lives ∪ m hellish lives would be a good population. Now, for every good population X and all bad populations Y and Z, if $X \cup Y$ would be good and Y is worse than Z, then $X \cup Z$ would be good. (To see this, assume that $X \cup Y$ would be good and that Y is worse than Z. Since Y is worse than Z, by Separability, $X \cup Z$ would be better than $X \cup Y$. And since $X \cup Y$ would be better than the empty population, by Transitivity, $X \cup Z$ would be better than the empty population – i.e. good.) Therefore, n heavenly lives ∪ any number of barely bad lives would be good.

Using a parallel strategy, we can show that (Revised) Heavy Tails similarly implies that there is a natural number m such that m hellish lives ∪ any natural number of barely good lives would be a bad population. The proof is omitted for brevity.

Calvin Baker is a PhD candidate in philosophy at Princeton University. He works on ethics, decision-making under normative uncertainty and Buddhist philosophy.

Cite this article: Baker C. Non-Archimedean population axiologies. *Economics and Philosophy*. https://doi.org/10.1017/S0266267124000099