

The Welfare Diffusion Objection to Prioritarianism

November 9, 2021

Abstract

According to the Welfare Diffusion Objection, Prioritarianism should be rejected because it implies the desirability of welfare diffusion: the implausible claim that it can be better for there to be less total wellbeing, spread thinly because a larger number of people exist, than for there to be more total wellbeing, spread more generously because a smaller number of people exist. Prioritarianism does not directly imply the desirability of welfare diffusion if it is construed as a theory of same-person comparisons only (as it is by Parfit 2003). This does not, however, mean that Prioritarians avoid the Welfare Diffusion Objection entirely. I argue that while Prioritarianism may not have any different-number *implications*, the Prioritarian is nevertheless implicitly committed to some fairly weak principles for comparing different-number populations. I show that these principles, when conjoined with a Prioritarian same-person axiology, imply the desirability of welfare diffusion. A related argument supports the view that populations should be ranked by their total non-priority-weighted wellbeing.

Keywords: *Prioritarianism, Population Ethics, Welfare Diffusion, Separability*

1 Introduction

According to Prioritarianism, “benefiting people matters more the worse off these people are” (Parfit 1997, p. 213). On the version of Prioritarianism I shall be concerned with,

- (i) Benefits to the worse off matter more in an *axiological* sense: they do more to make an outcome better than same-sized benefits to the better-off.
- (ii) The worse off are those who have less *wellbeing*, rather than (for example) those who have fewer resources.
- (iii) The wellbeing in question is overall *lifetime* wellbeing.
- (iv) Benefits to the worse off matter more only because these people are at a lower absolute level, and not because they are at a lower level relative to other people. (That is, Prioritarianism of the sort I am interested in rules out Egalitarianism.)

According to this version of Prioritarianism, it would be better for ten people to have slightly less than 50 units of wellbeing each – for concreteness, let’s say 49 units – than it would be for five to have 100, and the remaining five to have

nothing. This judgement is plausible. But what should a Prioritarian say about a similar case involving different numbers of people, in which we must choose between the existence of five people, each with 100 units of wellbeing, or the existence of ten people, each with 49 units of wellbeing? Perhaps Prioritarians should make the analogous judgement here: they should say that it would be better for there to be the greater number of individuals, even though each person who exists would be worse off. If so, then Prioritarians are, as Ingmar Persson puts it, committed to the *desirability of welfare diffusion* (2011, 2012). That would be a problem. The desirability of welfare diffusion is implausible: how could it be better for there to be more people, if this reduces the total amount of whatever makes life worth living? If Prioritarianism commits us to the desirability of welfare diffusion, it should be rejected. Or so goes the *Welfare Diffusion Objection* to Prioritarianism.

A number of authors have argued that Prioritarianism by itself implies the desirability of welfare diffusion.¹ The most obvious argument for this claim starts from the observation that Prioritarianism can be construed as recommending populations with greater total *priority-weighted* wellbeing, where the priority weighting function is some strictly concave, strictly increasing function of wellbeing which maps the neutral level to zero.² Since larger populations with less total wellbeing can nevertheless have greater *priority-weighted* wellbeing, Prioritarianism (allegedly) says that such populations are better (Persson 2011, 2012).

Alternatively, it might be argued that when we bring additional people into existence with lives worth living, these people are thereby benefited. Since Prioritarians have to say that benefiting people always makes the world better (according to this argument), they are committed to the premise that such additions are always a change for the better. As we shall see in §2, given Prioritarian judgements in same-person cases, and assuming transitivity, this premise implies the desirability of welfare diffusion.³

There are problems with both arguments. The first argument construes Prioritarianism as requiring different-number populations to be compared by their *sums* of priority-weighted wellbeing. But it is far from obvious that Prioritarianism must aggregate priority-weighted wellbeing levels by *summing* them, rather than by applying some other method of aggregation, such as taking the average. Until a case is made for aggregation by summation, the first argument is therefore incomplete.

The second argument takes Prioritarianism to be committed to the claim that the provision of benefits of any sort must always constitute a change for the better, if all else is equal. This argument is a little stronger, but it is not decisive. Consider the question of whether Existence Comparativism is true – that is, whether a good (bad) life can be better (worse) for a person than non-existence. This question is not yet settled.⁴ Since the question is not settled, it is open to the Prioritarian to deny Existence Comparativism. They could then claim that bringing additional lives worth living into existence does not

¹Persson 2011, Persson 2012, Holtug 2010.

²We shall discuss *Critical Level* versions of Prioritarianism, which do not map the neutral level to zero, in §3.

³By “same-person” cases, I mean cases involving the same people in each population under consideration. In contrast, “different-number” cases involve populations which have different numbers of people.

⁴Many philosophers reject Existence Comparativism. See for instance Narveson 1967, Parfit 1984, p. 489, Broome 1999, p. 168, Bykvist 2007 and Bader forthcoming.

constitute a provision of a benefit, since doing so does *not* make the additional people better off. Alternatively, the Prioritarian may simply stipulate that their view is not to be understood as including a commitment to the idea that the provision of a benefit constitutes a change for the better, when the benefit in question results from bringing an additional person into existence.⁵

More broadly, both arguments can be rejected on the basis that they mistakenly assume that Prioritarianism involves a commitment to *any* comparisons of different-number populations whatsoever. Derek Parfit, the most prominent advocate of Prioritarianism, thought that this way of thinking about Prioritarianism is mistaken. He claimed that Prioritarianism should be viewed as a theory of same-person comparisons *only*:

Like the Principles of Personal Good, or Pareto Principles, the Prioritarian Principles that I have considered cannot be applied to cases in which, in the different possible outcomes, different people would exist. When we consider these cases, we need other principles.

Parfit 2012, p. 440

So, is the Welfare Diffusion Objection toothless? I shall argue that it is not. If Prioritarianism is understood as a theory of comparisons of same-person populations, it cannot *imply* the desirability of welfare diffusion. Nevertheless, since Prioritarians are, as Parfit (1997, p. 214) puts it, concerned with people's absolute wellbeing levels rather than their wellbeing levels relative to other people, they are implicitly committed to some fairly minimal principles for comparing populations involving different numbers of people. These principles, conjoined with a Prioritarian axiology for same-person populations, imply the desirability of welfare diffusion.

Here's the plan for the paper. I begin in §2 by demonstrating that it is impossible to jointly satisfy a condition codifying avoidance of the desirability of welfare diffusion, a weak version of the Mere Addition Principle, and the Strong Pigou-Dalton Principle, which expresses the core Prioritarian idea that benefits to the worse off count for more. In §3, I argue that Prioritarians have plenty of reason to accept, and not much reason to reject, a weak version of the Mere Addition Principle. I also argue that Prioritarians are implicitly committed to the principle of *Separability*, which says roughly that the moral value of the existence of a person (or set of people) at a given wellbeing level (or distribution of wellbeing levels) is independent of the status of other people. Separability provides strong support for the weak version of the Mere Addition Principle. It provides even stronger support for a yet weaker version of the Mere Addition Principle, which is enough to imply the desirability of welfare diffusion if we assume a slightly stronger (but still compelling) version of Pigou-Dalton.

In §4, I consider a related argument for Totalism (the view on which one population is better than another iff it has greater total wellbeing). A weaker, more compelling version of the Strong Pigou-Dalton Principle of §2, the welfare diffusion avoidance condition and the weak version of the Mere Addition Principle turn out to imply Totalism, or rather its restriction to cases where nobody has a life worth not living. This argument is interesting both in its own right,

⁵Those who deny Existence Comparativism may nevertheless accept that existence with a life worth living can be *good* for a person, and that bringing someone into existence with such a life can constitute a *non-comparative* benefit to this person (McMahan 2013). Prioritarians who accept this view can still claim that Prioritarianism is to be understood as a theory only of the moral importance of comparative benefits.

and because it gives rise to an objection to the arguments of sections 2 and 3: these arguments assume (or come close to assuming) principles which jointly imply the Repugnant Conclusion. §5 replies to this objection, and to two other objections. I conclude in §6.

2 The Mere Addition Argument for Welfare Diffusion

We begin with a few technical preliminaries. As was implicit earlier, wellbeing levels are to be identified with real numbers. Positive numbers represent lives worth living, negative numbers represent lives worth not living, and zero represents a neutral life. Greater numbers represent better lives. There remains the question of how ratios of the differences between wellbeing levels are to be understood. I shall use the following scale: for some fixed good quality of life q , a life is at $x \geq 0$ units of wellbeing if and only if it is equally as good as a life consisting of x years of life, constantly at quality q . A life is at $-x$ if and only if one would rationally be indifferent between a 50-50 gamble yielding the $-x$ life and a life at x , or a neutral life for sure.⁶

I assume that there are infinitely many possible people. A *population* is an assignment of finitely many possible people to wellbeing levels. Populations represent the distributions in which precisely these people exist, with lives at the respective wellbeing levels, and nobody else exists. $p_i[w_i]$ denotes the population consisting of just person p_i at level w_i ; similarly, if X is a set of possible people, $X[w_i]$ denotes the population containing the X -people at level w_i , and nobody else. Populations are disjoint when they have no persons in common. When two populations X and Y are disjoint, we may write $X + Y$ for the population given by the union of the two assignments, which consists of the X -people at their respective levels, the Y -people at their respective levels, and nobody else.⁷

We shall be interested in the at-least-as-good-as relation, denoted by \succeq . We shall understand this to be a transitive binary relation on populations.⁸ This is not a completely innocent assumption. Some philosophers, most notably Larry Temkin (1987, 1996, 2012) and Stuart Rachels (1998, 2001, 2004), have argued that the at-least-as-good-as relation is not transitive. Others believe that the at-least-as-good relation is option-set-dependent, and hence cannot be understood to be a binary relation on populations.⁹ I shall set such positions aside until the conclusion of this paper.

That is enough in the way of background. The aim of the present section is to demonstrate that the core commitment of Prioritarianism, when conjoined with a weak version of the Mere Addition Principle, implies the desirability of welfare diffusion. We shall understand the core commitment of Prioritarianism

⁶Readers who doubt that additional years of good life have constant marginal value presumably have some other scale in mind; such readers are invited to consider that scale. I shall consider a different way of cardinalising wellbeing in §5.2.

⁷When writing such unions, the respective two populations are always assumed to be disjoint; quantifiers should be understood to be restricted accordingly, even if this is not explicitly mentioned.

⁸Recall that a binary relation R is *transitive* iff whenever aRb and bRc , we have aRc .

⁹See Frick (2014, forthcoming). Cusbert (2017) suggests that Temkin's Essentially Comparative View (2012) can be understood so as to imply option-set-dependent betterness, rather than intransitivity within choice sets.

to be the principle of

Strong Pigou-Dalton Let p_i and p_j be any two possible people, and let U be any disjoint unaffected background population. If w^- is a lower wellbeing level than w , then for any positive quantity of additional wellbeing a , there is a small positive quantity of wellbeing ϵ such that

$$U + p_i[w^- + a - \epsilon] + p_j[w] \succ U + p_i[w^- + \epsilon] + p_j[w + a]$$

Strong Pigou-Dalton just says that slightly smaller benefits to the worse off make the world better than slightly larger benefits to the better off. Prioritarians cannot reject this principle.

The weak version of the Mere Addition Principle is

Weak Mere Addition For any populations X and Y , if Y consists only of lives worth living, then $X + Y$ is not worse than X .

Weak Mere Addition does not imply that an addition of lives worth living is *at least as good*, or *better*, than no addition at all. Principles of that sort are suspect on the grounds that they conflict with the Evaluative Procreation Asymmetry, according to which bringing lives worth living into existence never makes the world better.¹⁰ The Weak Mere Addition Principle faces no such objection.

Finally, we shall need a principle which guarantees avoidance of the Welfare Diffusion Objection. This shall be:

Different-Number Egalitarian Dominance Let X and Y be any populations. If

- (i) X is a perfectly equal non-empty population of lives worth living;
- (ii) each person in X is better off than each person in Y ;
- (iii) each person in Y exists in X (and is therefore better off in X than in Y);
- (iv) X has greater total wellbeing than Y ,

then X is at least as good as Y .¹¹

These three principles are mutually inconsistent. That is, we have

Proposition 1 No population axiology satisfies Strong Pigou-Dalton, Weak Mere Addition and Different-Number Egalitarian Dominance.

Proof. Let p_i and p_j be any two possible people. Apply the Strong Pigou-Dalton principle with an empty unaffected background population, higher

¹⁰See for example Frick 2014, 2020, Roberts 2011. The best argument against the Evaluative Procreation Asymmetry that I know of is provided by John Broome (2004, 2005).

¹¹Note that in order for Different-Number Egalitarian Dominance to be false, it is enough for welfare diffusion to merely be *not undesirable* (not worse), rather than for it to be desirable (better). It seems to me, however, that the non-undesirability of welfare diffusion is not much more plausible than the desirability of welfare diffusion. I shall thus ignore this distinction going forward.

wellbeing level 50, lower wellbeing level 0, and 50 units of potential additional wellbeing. We then have that there is some small positive quantity of wellbeing ϵ such that

$$p_i[50] + p_j[50 - \epsilon] \succ p_i[100] + p_j[\epsilon]$$

Different-Number Egalitarian Dominance implies that

$$p_i[100] \succeq p_i[50] + p_j[50 - \epsilon]$$

Transitivity then implies that

$$p_i[100] \succ p_i[100] + p_j[\epsilon]$$

Which contradicts Weak Mere Addition. □

Here's the upshot. Since Prioritarians cannot deny Strong Pigou-Dalton, they must choose between Weak Mere Addition and Different-Number Egalitarian Dominance (given transitivity). In the next section, I shall argue that Prioritarians cannot reasonably reject Weak Mere Addition. I shall also argue that they are implicitly committed to a principle of Separability, which in turn provides strong support for Weak Mere Addition, and even stronger support for a weaker version of Mere Addition which turns out to be enough to imply the desirability of welfare diffusion if we somewhat strengthen our Pigou-Dalton principle.

3 Weak Mere Addition and Separability

What distinguishes Prioritarianism from Egalitarianism? Parfit (1997, p. 214) answers: while Egalitarians are concerned with relations between people's wellbeing levels and the wellbeing levels of others, Prioritarians are solely concerned with people's absolute wellbeing levels. Consider a situation in which one person is at 100 and another is at 0, when both could instead have been at 50. Egalitarians decry this situation because it involves inequality: the less well-off person is not as well off as the better-off person. Prioritarians decry the same situation on different grounds: for them, the difference between 0 and 50 units of wellbeing is morally more important than the difference between 50 and 100 units of wellbeing, even though the wellbeing differences are of the same size.

A sole concern with people's absolute wellbeing levels provides immediate support for the Weak Mere Addition Principle. Consider its application in the proof of Proposition 1. It there implied that $p_i[10] + p_j[\epsilon]$ is not worse than $p_i[10]$, where ϵ is some arbitrarily small positive quantity of wellbeing. Egalitarians can reject this claim. They can say that, because the existence of p_j introduces inequality, it would be better if only p_i were to exist. This move is unavailable to Prioritarians. Since they are concerned only with people's absolute levels of wellbeing, they cannot appeal to relations between p_i 's wellbeing and p_j 's wellbeing when both exist. On the face of it, it could only be bad for p_j to exist on Prioritarian grounds if existence is bad *for* p_j . But that is false: while p_j has only a low positive wellbeing level, a low positive level is still positive, and we stipulated in §2 that a wellbeing level is represented by a positive number if and only if lives at that level are worth living.

The argument from Prioritarians' sole concern with absolute levels to the Weak Mere Addition Principle can be tightened up. A sole concern for people's absolute wellbeing levels is captured by the principle of

Separability Let X, Y and Z be any populations. X is at least as good as Y if and only if $X + Z$ is at least as good as $Y + Z$.

Separability is widely accepted by Prioritarians. Indeed, Adler & Holtug (2019, p. 104) take a version of Separability to be part of the *definition* of Prioritarianism. The version of Separability they are talking about is restricted to same-person cases, while mine is unrestricted (and needs to be).¹² But a sole concern for people’s absolute wellbeing levels supports the unrestricted version of Separability just as well as it supports the restricted version. If (unrestricted) Separability is false, then the relative contributive values of populations X and Z depend not only on the absolute levels of the persons involved in X and Y , but also on the status of the unaffected people in population Z . Thus, one cannot deny (unrestricted) Separability without thereby expressing a concern for more than just people’s absolute wellbeing levels.

Separability is hard to square with the negation of Weak Mere Addition. If we deny Weak Mere Addition, we think that sometimes it is worse to add lives worth living to the world. If we accept Separability as well, then we think that it is *always* worse to add such lives to the world. Apart from its prima facie implausibility, this claim turns out to force us to deny the compelling

*Absolute Value Principle*¹³ If X is a population consisting solely of lives worth living, and Y is a population consisting solely of lives worth not living, then X is better than Y .

At least, this is so if we accept

Non-Absolute Priority For any positive quantity of wellbeing x , there is some sufficiently small positive quantity of wellbeing ϵ such that for any persons p_i and p_j , and any disjoint unaffected background population U ,

$$U + p_i[\epsilon] + p_j[\epsilon] \not\prec U + p_i[x] + p_j[-\epsilon]$$

Non-Absolute Priority says that it would be better to benefit someone who is better off, rather than to benefit someone who is badly off by some miniscule amount, even if that amount pushes them from a life barely worth not living to a life worth living. The opposite view, Absolute Prioritarianism, says that those who are below the threshold of a life worth living are to be prioritised absolutely over those who are above the threshold. On this view, it would be better to spare one person from a pinprick which would push them just barely below the neutral level, rather than to spare trillions of people from a greater harm which would not push them below the neutral level. Since most people do not find this kind of view very plausible, I shall not discuss it further.¹⁴

We return to the incompatibility between the Absolute Value Principle and the negation of Weak Mere Addition. If Weak Mere Addition is false, there is some case in which an addition of a life at positive wellbeing level x , say of

¹²Many thanks to an anonymous referee for pointing this out.

¹³This principle is often called “Priority for Lives Worth Living” (see for instance Blackorby et al. 2005, p. 135). I avoid this name because it is suggestive of Prioritarianism, whereas the Absolute Value Principle is satisfied by many non-Prioritarian population axiologies (such as Total and Average Utilitarianism).

¹⁴But see Crisp 2003 for critical discussion of this kind of Absolute Prioritarianism.

person p_2 , is worse than no addition at all.¹⁵ By Separability, adding p_2 at level x is therefore *always* worse than no adding no one. Now consider the following three populations.

$$\begin{aligned} A_1 & p_1[-\epsilon] \\ B_1 & p_1[-\epsilon] + p_2[x] \\ C_1 & p_1[\epsilon] + p_2[\epsilon] \end{aligned}$$

By the negation of Weak Mere Addition and Separability, B_1 is worse than A_1 . By the Absolute Value Principle, A_1 is worse than C_1 . Transitivity then implies that B_1 is worse than C_1 , contradicting Non-Absolute Priority.

This argument for Weak Mere Addition seems to me decisive, but not everyone accepts the Absolute Value Principle. Critical Level Prioritarians, for example, believe that it can be worse for there to be many lives that are positive, but below some critical level $x^* > 0$, rather than for there to be fewer lives at a negative wellbeing level.¹⁶ So let me provide another sort of argument. This argument needs a different version of Non-Absolute Priority and a stronger Pigou-Dalton principle, but in return weakens the Absolute Value Principle to the

Weak Absolute Value Principle There is some positive wellbeing level a such that if X is a population consisting solely of lives which are at least at wellbeing level a , and Y is a population consisting solely of bad lives, then X is better than Y .

Critical Level Prioritarians deny the Absolute Value Principle by saying that a large number of lives barely worth living can be worse than a smaller number of lives worth not living. But they do not say that a large number of *excellent* lives can be worse than a smaller number of negative lives. The former is merely implausible; the latter is *absurd*.

The non-absolute priority condition we need is

Non-Absolute Priority 2 For some sufficiently small positive quantity of wellbeing ϵ' , if \mathcal{W} is any bounded interval of wellbeing levels, there is a sufficiently large positive quantity of wellbeing δ' such that given any unaffected background population U , if $\epsilon \leq \epsilon'$, $\delta \geq \delta'$ and w_i, w_j are in \mathcal{W} , then

$$U + p_i[w_i] + p_j[w_j] \prec U + p_i[w_i - \epsilon] + p_j[w_j + \delta]$$

This principle looks a little complicated, but essentially it just says that it is always better to provide a sufficiently large benefit to a better-off person, rather than a very small benefit to a worse-off person. The required size of the benefit to the better-off person can increase as the gap between the two increases. We can also restrict the principle to cases in which the worse-off person has a life which is, at worst, only *slightly* bad for them.

Consider now the following three populations, where a is a sufficiently high wellbeing level for the Weak Absolute Value Principle to apply:

¹⁵Strictly speaking, this does not follow from the negation of Weak Mere Addition, since Weak Mere Addition could be false because some addition of *multiple* lives worth living is worse, while additions of individual lives worth living are always incomparable with no addition at all. In practice this does not matter, because the instance of Weak Mere Addition appealed to in the proof of Proposition 1 concerned an addition of a single life.

¹⁶See Blackorby et al. 1995, 2005 for a discussion of critical level views.

$$\begin{aligned}
A_2 & p_i[-\epsilon] \\
B_2 & p_i[-\epsilon] + p_j[a^+] \\
C_2 & p_i[a] + p_j[a]
\end{aligned}$$

By applying Non-Absolute Priority 2 finitely many times, together with transitivity, it can be shown that B_2 is better than C_2 .¹⁷ The Weak Absolute Value Principle implies that C_2 is better than A_2 . We thus have that B_2 is better than A_2 . We only really need the weaker conclusion that B_2 is not worse than A_2 . Given Separability, we then have

Very Weak Mere Addition There is some positive wellbeing level a such that for any population X , and any population Y consisting of lives at level a , $X + Y$ is not worse than X .

To complete our argument, we shall need the following stronger version of Pigou-Dalton:

Stronger Pigou-Dalton For any positive wellbeing level a , there exists a wellbeing level $a^+ > 2a$ and a small positive quantity of wellbeing ϵ such that for any possible people p_i and p_j , and any unaffected background population U ,

$$U + p_i[a^+/2 - \epsilon] + p_j[a^+/2 - \epsilon] \succ U + p_i[a^+] + p_j[a]$$

Stronger Pigou-Dalton says that it is sometimes better to spread wellbeing evenly between two people, even when this comes at the cost of slightly more than a units of total wellbeing. For example, it is very plausible that it would be better for two people to be at level 499 than it would be for one of these people to be at level 1000 and the other to be at level 100. (This example sets $a^+ = 1000$, $a = 100$, $\epsilon = 1$.) We shall only need to apply Stronger Pigou-Dalton once, for some value of a for which the Weak Absolute Value Principle holds (such as $a = 100$), so it would be sufficient for the Prioritarian to accept the plausible judgement just in this example, even if they do not accept Stronger Pigou-Dalton in general.¹⁸

It can now be shown that

Proposition 2 No population axiology satisfies Stronger Pigou-Dalton, Very Weak Mere Addition and Different-Number Egalitarian Dominance.

¹⁷Let e' be a small enough quantity of wellbeing that Non-Absolute Priority 2 applies. Let n be the smallest number greater than $\frac{a+\epsilon}{e'}$, and let $e = \frac{a+\epsilon}{n}$; we then have $e \leq e'$. We can apply Non-Absolute Priority 2 repeatedly to show that

$$p_i[a] + p_j[a] \prec p_i[a - e] + p_j[a + \delta_1] \prec \dots \prec p_i[a - ne] + p_j[a + \sum_{i=1}^n \delta_i]$$

Writing a^+ to stand for $a + \sum_{i=1}^n \delta_i$, the last population is equal to

$$p_i[-\epsilon] + p_j[a^+].$$

¹⁸Why should we accept that the Weak Absolute Value Principle holds for $a = 100$? Recall that our wellbeing scale corresponds to “years of good quality life”. It seems to me that any population of negative lives is clearly worse than a population of one-hundred-year-long lives at a constant good quality.

Proof. Let a be a wellbeing level witnessing Very Weak Mere Addition (that is, additions at level a are never worse). Stronger Pigou-Dalton implies that there exist $a^+ > 2a$ and a small positive quantity of wellbeing ϵ such that

$$p_i[a^+/2 - \epsilon] + p_j[a^+/2 - \epsilon] \succ p_i[a^+] + p_j[a]$$

Different-Number Egalitarian Dominance implies that

$$p_i[a^+] \succeq p_i[a^+/2 - \epsilon] + p_j[a^+/2 - \epsilon]$$

Applying transitivity, we then conclude that

$$p_i[a^+] \succ p_i[a^+] + p_j[a]$$

However, this last claim contradicts Very Weak Mere Addition (since a is a witness to that principle.) \square

Every plausible extension of Prioritarianism to different-number cases, including Critical Level Prioritarianism, satisfies the Weak Absolute Value Principle. More tentatively, it seems to me that every plausible version of Prioritarianism satisfies Non-Absolute Priority and Stronger Pigou-Dalton, or at least satisfies the latter principle in *some* cases, like the one I outlined. If all that is right, then every plausible extension of Prioritarianism to different-number cases satisfying Separability violates Different-Number Egalitarian Dominance, and is thereby open to the Welfare Diffusion Objection.

It follows that if Prioritarianism involves an implicit commitment to unrestricted Separability (as I have argued), it is open to the Welfare Diffusion Objection. And even Prioritarians who deny Separability do not necessarily escape the Welfare Diffusion Objection. Weak Mere Addition is independently plausible, and Very Weak Mere Addition is even more plausible. Proposition 2 shows that Prioritarians who accept Very Weak Mere Addition and Stronger Pigou-Dalton cannot escape the Welfare Diffusion Objection, even if they do not accept Separability.¹⁹

4 A Related Argument for Totalism

The arguments of §2 and §3 are not only of interest to Prioritarians and their critics. Proposition 1, which we used to establish that Prioritarians cannot accept Weak Mere Addition without leaving themselves open to the Welfare Diffusion Objection, can be repurposed into an argument for Totalism. Recall that Proposition 1 shows that no population axiology satisfies all of Weak Mere Addition, Different-Number Egalitarian Dominance, and Strong Pigou-Dalton. In so far as Weak Mere Addition and Different-Number Egalitarian Dominance are compelling, we thus have a compelling argument for the negation of Strong Pigou-Dalton. The argument can be tightened up further. Strong Pigou-Dalton is controversial (albeit still intuitively compelling) because it says that benefits to the worse off matter *more* than benefits to the better off. A related principle is at least as intuitively compelling as Strong Pigou-Dalton without being controversial in the same way, namely

¹⁹Or, as we have seen, it is enough for the Prioritarian to accept the verdict of Stronger Pigou-Dalton for some value of a which plausibly witnesses Very Weak Mere Addition.

Weak Pigou-Dalton Let p_i and p_j be any two possible people, and let U be any disjoint unaffected background population. If w^- is a lower well-being level than w , then for any positive quantity of additional wellbeing a ,

$$U + p_i[w^- + a] + p_j[w] \succeq U + p_i[w^-] + p_j[w + a]$$

Assume also the principle of Weak Mere Addition* (which is exactly like Weak Mere Addition, except that it also rules out that additions of neutral lives can be worse). Finally, we need a slightly stronger version of Different-Number Egalitarian Dominance, which applies to neutral as well as good lives and drops the requirement that the people in the (possibly) smaller population must exist in the (possibly) larger population:

*Different-Number Egalitarian Dominance** Let X and Y be any populations. If

- (i) X is a perfectly equal non-empty population of good or neutral lives;
- (ii) each person in X is at least as well off as each person in Y ;
- (iii) X has greater total wellbeing than Y ,

then X is at least as good as Y .

We now have all the ingredients needed to argue for

Totalism for Good Populations Suppose that non-empty populations X and Y contain only lives that are neutral or good. Then X is at least as good as Y if and only if $T(X)$ is at least as great as $T(Y)$ (where $T(Z)$ denotes the total wellbeing of population Z).

That is, we have

Proposition 3 Every population axiology which satisfies Weak Mere Addition*, Weak Pigou-Dalton and Different-Number Egalitarian Dominance* also satisfies Totalism for Good Populations.²⁰

Proof. Given transitivity, it is sufficient to show that every population consisting only of good or neutral lives is equal in value to a singleton population containing one person at the total wellbeing level. Different-Number Egalitarian Dominance* then requires these singleton populations to be ranked according to total wellbeing, and transitivity extends this ranking to all other populations with only good or neutral lives.

Let C be an arbitrary non-empty population of good or neutral lives. Let p_i be some person in C , and let C' be the set of people in C , except for p_i . Define populations A and B to be:

$$\begin{aligned} A & p_i[T(C)] \\ B & p_i[T(C)] + C'[0] \end{aligned}$$

²⁰Huemer (2012) provides a similar argument, although this argument assumes a stronger version of the Mere Addition Principle, which is justified by appealing to the controversial principle of Existence Comparativism.

C is obtainable from B by means of a series of transfers of wellbeing from better-off to worse-off (taking wellbeing from p_i each time). Weak Pigou-Dalton therefore implies that C is at least as good as B . Different-Number Egalitarian Dominance* implies that A is at least as good as C . Applying transitivity, we find that A is at least as good as B . But Weak Mere Addition* implies that B is not worse than A . It follows that A and B must be equally good.²¹ Recalling that A is at least as good as C and that C is at least as good as B , we can conclude that A and C are equally good. \square

5 Objections and Replies

5.1 The Repugnant Conclusion

Proposition 2 strikes me as a sound argument for Totalism for Good Populations. Many readers will presumably disagree with this assessment, since Totalism for Good Populations implies the Repugnant Conclusion, and most philosophers are not prepared to accept the Repugnant Conclusion.²² One might worry that this makes the arguments of §2 and §3 suspect: these arguments assume (or near enough assume) principles which jointly entail the Repugnant Conclusion.²³

I do not think that a reply on these lines is of much help to the Prioritarian. Consider first Weak Mere Addition. This principle, while independently plausible, is made more plausible by the Prioritarian focus on people's absolute wellbeing levels. If the best way to avoid the Repugnant Conclusion is to reject Weak Mere Addition, Prioritarians are in a tough spot.

If, on the other hand, the way to avoid the Repugnant Conclusion is via denial of Weak Pigou-Dalton, the Prioritarian is again in trouble, for it is hard to see how a Prioritarian *can* reject Weak Pigou-Dalton. Once again: so much the worse for Prioritarianism. Even if Prioritarians could somehow reject Weak Pigou-Dalton in order to avoid the Repugnant Conclusion, it is far from clear that they can do so in a way that will avoid the Welfare Diffusion Objection. Suppose for instance that a Prioritarian can adopt something like Parfit's (2004) Perfectionism. On this theory, pure transfers of wellbeing from better-off to worse-off can make the world worse when they involve the loss of the best things in life, but not when they do not. Since we can use Pigou-Dalton principles to construct violations of Different-Number Dominance which do not involve losses of the best things in life, Perfectionism will not help Prioritarians avoid the Welfare Diffusion Objection, even though it is inconsistent with Pigou-Dalton principles. For example, Prioritarians could likely be pushed to

²¹We have $A \succeq B$ and $B \not\succeq A$. Since (by definition) $B \prec A$ iff $A \succeq B$ and $B \not\succeq A$, we have that $B \succeq A$, hence $A \sim B$.

²²According to the Repugnant Conclusion, for any possible population of at least ten billion people, all with a very high quality of life, there must be some much larger imaginable population whose existence, if other things are equal, would be better, even though its members have lives that are barely worth living (Parfit 1984, p. 388). Although the Repugnant Conclusion has traditionally been regarded as a decisive counterexample to Totalism and other population axiologies, this traditional view no longer enjoys near-unanimity; see Zuber et al. 2021.

²³Different-Number Egalitarian Dominance is a premise of Proposition 1. Weak Mere Addition* is almost identical to Weak Mere Addition, which is a premise of Proposition 1. Weak Pigou-Dalton is closely related to Strong Pigou-Dalton, which is the last premise of Proposition 1.

accept that a population consisting of ten billion people, each at level 100, would not be better than a population of eleven billion people, each at level 90.²⁴ The latter population could retain the best things in life, since 90 years each of good quality life provides lots of space for excellence. But the claim that the latter population is not better than the former still violates Different-Number Egalitarian Dominance, and still leaves Prioritarians open to the Welfare Diffusion Objection.

The third possibility is that the Repugnant Conclusion is to be avoided by denying Different-Number Egalitarian Dominance*. But this idea is, on face of it, wrong-headed. The point of Different-Number Egalitarian Dominance* is to say that smaller populations are better than larger ones when each person in the smaller population is better off than each person in the larger population, provided that the smaller population *also* has greater total wellbeing. Those who wish to avoid the Repugnant Conclusion will presumably accept this claim, along with the stronger claim that the smaller population is often better even if it has *less* total wellbeing.

5.2 The Cardinalisation Objection

I have argued that Prioritarianism is open to the Welfare Diffusion Objection. But does the Welfare Diffusion Objection really have any bite? It might not, depending on how we generate the scale of wellbeing. I have generally worked with a scale based on years of good life: one unit of wellbeing corresponds to one year of good quality life. But we could generate the scale of wellbeing in another way: by appealing to utility functions generated by Expected Utility Theory (on the assumption that the axioms of Expected Utility Theory are satisfied for the prudential betterness relation on prospects).²⁵ If our wellbeing scale is generated in this way, it is at best unclear that we can have sensible intuitions regarding total amounts of wellbeing.²⁶ The reason is that we plausibly lack an intuitive grasp of the utility functions from which our wellbeing scale is generated, since these utility functions are mere representation devices, spit out by a complex mathematical theorem. If our intuitions regarding total quantities of wellbeing are baseless, then it seems the Welfare Diffusion Objection is nothing to worry about, since it appeals to precisely these sorts of baseless intuitions. Call this the “Cardinalisation Objection”.

I do not think that the Cardinalisation Objection is of much help to Prioritarians. The reason is that it works just as well against Prioritarianism itself: we have no reason to accept that same-sized benefits to the worse-off matter more if we lack sensible intuitions concerning the moral importance of quantities of wellbeing. The Cardinalisation Objection can therefore only render the Welfare Diffusion Objection impotent by rendering it unnecessary.

One can also respond to the Cardinalisation Objection by ignoring it. Whatever the merits of cardinalising wellbeing via Expected Utility Theory, we can

²⁴Begin with a population A of ten billion at 100. Add one billion at level 1, and raise the A -people to level 101, resulting in population B . Finally, let C be all eleven billion people at 90. The Prioritarian will presumably judge that C is better than B (or, if not, the numbers may be adjusted as necessary). Weak Mere Addition implies that B is not worse than A . Transitivity then yields that C is not worse than A , contradicting Different-Number Egalitarian Dominance.

²⁵See Morgenstern & Von Neumann 1944, Savage 1954 or Fishburn 1982, among many others.

²⁶This line of thought is suggested by Greaves 2015.

still talk about other wellbeing scales for which our intuitions regarding quantities of wellbeing are not baseless, such as the “years of good life” scale. At least on the face of it, Prioritarians are committed to Strong Pigou-Dalton and Different-Number Egalitarian Dominance on the years-of-good-life scale. So the Welfare Diffusion Objection seems to be a live problem for Prioritarianism on the years-of-good-life scale, even if it is not a problem on the scale generated by Expected Utility Theory.

5.3 Quarantine

Yet another response to the Welfare Diffusion Objection is to attempt to “quarantine” the problem, along with all other difficulties associated with different-number cases. Parfit summarises this strategy with an analogy:

It’s very difficult to formulate acceptable welfarist theories that could apply to cases that involve infinite quantities of such things as suffering and happiness. That’s a worry, but it doesn’t undermine our confidence in the theories that can handle cases with only finite quantities.

Parfit 2004, p. 257

Similarly, one might think that although it is difficult to formulate acceptable theories of different-number comparisons, this should not undermine our confidence in our theories of same-person comparisons.

It seems to me that there is a more promising and a less promising interpretation of this strategy. On the more promising interpretation, when *all* theories applicable to domain D face severe difficulties when extended to the larger domain D' , in some cases this should not decrease our confidence in the theories applicable to D . On the less promising interpretation, if theory T applicable to domain D faces some particular difficulty whenever it is extended to domain D' , this should not make us sceptical of T , even if some other theory T' applicable to D can be extended to D' without facing a similar difficulty.

Since some non-Prioritarian views *can* avoid the Welfare Diffusion Objection when extended to different-number cases (for example, Totalism), only the less promising interpretation of the strategy will help to defend Prioritarianism. Yet it seems to me that Parfit’s analogy does not help to make the strategy seem plausible on this interpretation. Consider two theories, T and T' , applicable to the evaluation of populations involving finite quantities of suffering and happiness. If some extension of T can deal with populations involving infinite quantities of suffering and happiness in an acceptable way, and no extension of T' deals with such populations in an acceptable way, this does seem to me to provide strong reason to prefer T over T' .

6 Conclusion

I have argued that the Welfare Diffusion Objection poses a significant challenge to Prioritarianism, even on the assumption that Prioritarianism does not directly imply the desirability of welfare diffusion. Importantly, I have assumed throughout this paper that the at-least-as-good-as relation is transitive, and that it is independent of the option set in which the relevant two populations are to be compared. It seems to me that if Prioritarians want to avoid the

Welfare Diffusion Objection, their best bet may be to do so by challenging one or the other of these assumptions. Taking such a path would represent a significant departure from the sort of traditional population axiology which has been the focus of this paper, and would open the Prioritarian to significant pragmatic challenges – especially, the prospect of susceptibility to value pumps.²⁷ Still, perhaps these and other objections can be answered.²⁸

Otherwise, there are several possible responses to the arguments of this paper. One is to simply bite the bullet and accept the desirability of welfare diffusion. While this position seems to me unattractive, it could fairly be said that every transitive population axiology takes one unattractive position or another. The desirability of welfare diffusion is not clearly more implausible than other controversial positions in population axiology, such as acceptance of the Repugnant Conclusion.²⁹ It also may be that adopting a version of Prioritarianism which implies the desirability of welfare diffusion has payoffs elsewhere. For instance, Total Prioritarianism, unlike Totalism, implies that it is better to avoid the creation of people with lives at wellbeing level $-x$ (where x is positive), rather than to ensure the existence of people with lives at wellbeing level x (Holtug 2010, p. 255).³⁰ Total Prioritarianism also (again, unlike Totalism) implies, rather plausibly, that it can be better for there to be more total negative wellbeing spread thinly among a larger number of people than for there to be less total negative wellbeing spread thickly among a smaller number of people (Holtug 2010, pp. 256–257).

How bad would it be for a Prioritarian to admit welfare diffusion? The precise answer to this question depends on the extent to which priority is given to the worse-off, but let me give an indication of what might be true. It is plausible that it is always better to bring one person from wellbeing level 1 to level 40 than it is to bring a different person from level 40 to level 100. Prioritarians who accept this claim, along with Weak Mere Addition, will have to accept that it would not be worse for there to be twenty billion people with 40 years of good life, rather than ten billion people with 100 years of good life. If we consider cases where more people are added, the results can become intuitively more troubling. Suppose that it is always better to bring two people from level 1 to level 20 than it is to bring one person from level 20 to level 100. If that is true, and Weak Mere Addition is also true, it would not be worse for there to be thirty billion people at level 20 than it would be for there to be ten billion people at level 100. These conclusions are difficult, but perhaps not impossible, to accept.

Another option is to deny the Weak Mere Addition principle. As I argued extensively in §3, this option should not be taken by a Prioritarian. But one could abandon Prioritarianism for this reason, and instead accept Egalitarian-

²⁷A particularly compelling money/value pump for cyclic theories has recently been provided by Gustafsson & Rabinowicz (2020). Gustafsson (n.d.) further claims that intransitive, acyclic theories are also vulnerable to value pumps, although this argument is less secure than the value pump argument against cyclicity. Value pumps would, on the face of it, appear to be effective against option-set-dependent theories which involved cycles of betterness among pairwise choices.

²⁸Perhaps by defending a decision theory involving resolute choice (McClennen 1985), or to some other unorthodox decision theory (Ahmed 2017).

²⁹However, the “Super-Repugnant Conclusion” (Holtug 2010, ch. 9), which might be difficult to avoid for a Prioritarian who accepts the desirability of welfare diffusion, is more implausible than the Repugnant Conclusion.

³⁰By “Total Prioritarianism”, I mean the view that ranks populations according to their total priority-weighted wellbeing.

ism. This seems to me a reasonable response to the arguments of this paper: it is not crazy to claim that it can be worse to add people with lives worth living when (and because) doing so would introduce significant inequality. However, Proposition 2 shows that this position will need to involve the rejection of Very Weak Mere Addition, which may be difficult even for an Egalitarian.

A final option is to accept all of the premises of Proposition 3, taking the desirability of welfare diffusion to rule out Prioritarianism, and denial of Weak Mere Addition to rule out Egalitarianism. One would then be left with Totalism for Good Populations. The obvious next step is to accept unrestricted Totalism, but it's worth noting that one is not forced to this position. It is compatible with the premises of Proposition 3 to believe that priority is to be given to the worse off only when the worse off have lives worth not living. There is more to be said for this position than it might seem. Roger Crisp (2003, p. 755), noting the apparent absurdity of prioritising the rich over the super-rich, claims it is plausible that “when people reach a certain level, even if they are worse off than others, benefiting them does not, in itself, matter more”. If he is right, then there is some threshold after which considerations of priority no longer apply. If we think such a threshold should be non-arbitrary, what better candidate could there be than the neutral level of wellbeing?

References

- Adler, M. D. & Holtug, N. (2019), ‘Prioritarianism: A response to critics’, *Politics, Philosophy & Economics* **18**(2), 101–144.
- Ahmed, A. (2017), ‘Exploiting cyclic preference’, *Mind* **126**(504), 975–1022.
- Bader, R. M. (forthcoming), Person-affecting utilitarianism, in G. Arrhenius, K. Bykvist & T. Campbell, eds, ‘Oxford Handbook of Population Ethics’, Oxford University Press, Oxford.
- Blackorby, C., Bossert, W. & Donaldson, D. (1995), ‘Intertemporal population ethics: Critical-level utilitarian principles’, *Econometrica* **63**(6), 1303–1320.
- Blackorby, C., Bossert, W. & Donaldson, D. (2005), *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*, Cambridge University Press, New York.
- Broome, J. (1999), *Ethics Out Of Economics*, Cambridge University Press, Cambridge.
- Broome, J. (2004), *Weighing Lives*, Oxford University Press, Oxford.
- Broome, J. (2005), ‘Should we value population?’, *Journal of Political Philosophy* **13**(4), 399–413.
- Bykvist, K. (2007), ‘The benefits of coming into existence’, *Philosophical Studies* **135**(3), 335–362.
- Crisp, R. (2003), ‘Equality, priority, and compassion’, *Ethics* **113**(4), 745–763.
- Cusbert, J. (2017), ‘Acting on essentially comparative goodness’, *Thought: A Journal of Philosophy* **6**(2), 73–83.
- Fishburn, P. C. (1982), *The Foundations of Expected Utility*, Reidel, Dordrecht.

- Frick, J. (2014), ‘Making People Happy, Not Making Happy People’: A Defense of the Asymmetry Intuition in Population Ethics, PhD thesis, Harvard University.
- Frick, J. (2020), ‘Conditional reasons and the procreation asymmetry’, *Philosophical Perspectives: Ethics* **34**, 53–87.
- Frick, J. (forthcoming), Context-dependent betterness and the mere addition paradox, in J. McMahan, T. Campbell, J. Goodrich & K. Ramakrishnan, eds, ‘Ethics and Existence: the Legacy of Derek Parfit’, Oxford University Press, Oxford.
- Greaves, H. (2015), ‘Antiprioritarianism’, *Utilitas* **27**(1), 1–42.
- Gustafsson, J. E. (n.d.), Money-pump arguments. Unpublished manuscript.
- Gustafsson, J. E. & Rabinowicz, W. (2020), ‘A simpler, more compelling money pump with foresight’, *The Journal of Philosophy* **117**(10), 578–589.
- Holtug, N. (2010), *Persons, Interests, and Justice*, Oxford University Press, Oxford.
- Huemer, M. (2012), ‘Against equality and priority’, *Utilitas* **24**(4), 483–501.
- McClellenn, E. F. (1985), Prisoner’s dilemma and resolute choice, in R. Campbell & L. Sowden, eds, ‘Paradoxes of Rationality and Cooperation: Prisoner’s Dilemma and Newcomb’s Problem’, University of British Columbia Press, Vancouver, chapter 5, pp. 94–104.
- McMahan, J. (2013), ‘Causing people to exist and saving people’s lives’, *The Journal of Ethics* **17**(1), 5–35.
- Morgenstern, O. & Von Neumann, J. (1944), *Theory of Games and Economic Behavior*, Princeton University Press, Princeton.
- Narveson, J. (1967), ‘Utilitarianism and new generations’, *Mind* **76**(301), 62–72.
- Parfit, D. (1984), *Reasons and Persons*, Clarendon Press, Oxford.
- Parfit, D. (1997), ‘Equality and priority’, *Ratio* **10**(3), 202–221.
- Parfit, D. (2003), ‘Justifiability to each person’, *Ratio* **16**(4), 368–390.
- Parfit, D. (2004), Overpopulation and the quality of life, in J. Ryberg & T. Tännsjö, eds, ‘The Repugnant Conclusion: Essays on Population Ethics’, Kluwer Academic, London, pp. 145–164.
- Parfit, D. (2012), ‘Another defence of the priority view’, *Utilitas* **24**(3), 399–440.
- Persson, I. (2011), ‘Prioritarianism, levelling down and welfare diffusion’, *Ethical Theory and Moral Practice* **14**(3), 307–311.
- Persson, I. (2012), ‘Prioritarianism and welfare reductions’, *Journal of Applied Philosophy* **29**(4), 289–301.
- Rachels, S. (1998), ‘Counterexamples to the transitivity of *Better Than*’, *Australasian Journal of Philosophy* **76**(1), 71–83.

- Rachels, S. (2001), ‘A set of solutions to parfit’s problems’, *Noûs* **35**(2), 214–238.
- Rachels, S. (2004), Repugnance or intransitivity: A repugnant but forced choice, *in* J. Ryberg & T. Tännsjö, eds, ‘The Repugnant Conclusion: Essays on Population Ethics’, Kluwer Academic, London, pp. 163–186.
- Roberts, M. A. (2011), ‘The Asymmetry: A Solution’, *Theoria* **77**(4), 333–367.
- Savage, L. J. (1954), *The Foundations of Statistics*, John Wiley & Sons, New York.
- Temkin, L. S. (1987), ‘Intransitivity and the mere addition paradox’, *Philosophy & Public Affairs* **16**(2), 138–187.
- Temkin, L. S. (1996), ‘A continuum argument for intransitivity’, *Philosophy & Public Affairs* **25**(3), 175–210.
- Temkin, L. S. (2012), *Rethinking the Good: Moral Ideals and the Nature of Practical Reasoning*, Oxford University Press, Oxford.
- Zuber, S. et al. (2021), ‘What should we agree on about the repugnant conclusion?’, *Utilitas* .