

Against Equality and Priority

MICHAEL HUEMER
University of Colorado

I start from three premises, roughly as follows: (1) that if possible world x is better than world y for every individual who exists in either world, then x is better than y ; (2) that if x has a higher average utility, a higher total utility, and no more inequality than y , then x is better than y ; (3) that *better than* is transitive. From these premises, it follows that equality lacks intrinsic value, and that benefits given to the worse-off contribute no more to the world's value than equal-sized benefits given to the better-off.

I. THE DEBATE OVER EGALITARIANISM AND THE PRIORITY VIEW

Consider a world in which some are much better off than others. Suppose it is possible to benefit those who are worse-off, at the expense of the better-off. The better-off would lose a little more utility than the worse-off would gain, so the world's total utility would be slightly lowered; however, the change would bring about perfect equality. Assume this change could be effected without violating anyone's rights or otherwise committing any injustice (for instance, the better-off could be persuaded voluntarily to donate resources to the worse-off), and assume no other evaluatively relevant effects would occur. In this case, many believe that the proposed change would be for the better.

There are at least two reasons why this might be true. The first is the *Egalitarian* reason: we may believe that inequality is intrinsically bad.¹ Though the mentioned change would slightly lower the total utility of the world, it would also remove a large inequality. The reduction of inequality would be a good that would outweigh the loss in utility, leaving the world overall better.

This view is exposed to the popular *Levelling Down Objection*: if we hold that inequality is intrinsically bad, then we must also hold that merely harming the better-off is good in one respect. Imagine a second scenario: some people are much better off than others, but unfortunately in this case it is impossible to help the disadvantaged. It is, however, possible to harm the better-off; if we do so, everyone will then be at the same, low level of utility. If inequality is intrinsically bad, then this change would be good *in one respect* – that of eliminating inequality – even if it benefited no one. Indeed, it would be good in

¹ Larry Temkin, 'Inequality', *Philosophy and Public Affairs* 15 (1986), pp. 99–121, at 100; Kai Nielsen, *Equality and Liberty* (Totowa, NJ, 1985), p. 283.

that respect even if the change reduced everyone to a welfare level below the level the worse-off group enjoyed before the change. The Egalitarian need not embrace the extremely counter-intuitive claim that such a change would be good *overall*, but he must accept that a harm to everyone might be good in one respect. To some, this seems an unacceptably perverse conclusion.

This leads some to prefer the second reason for taking the redistribution of advantages discussed in the opening paragraph to be good. According to the *Priority View*, the redistribution is desirable overall, not because there is anything good in harming the advantaged, but simply because helping the disadvantaged is *more important* than maintaining the welfare of the presently advantaged. Everyone's well-being counts, every advantage adds to the goodness of the world, but an advantage of a given size counts *more* when given to someone whose initial welfare level is lower. Many find this view intuitively plausible.² The view avoids the Levelling Down Objection, and it avoids any appearance of perverseness, since proponents of the Priority View do not value harming anyone in any way.

This argument for preferring the Priority View to Egalitarianism has been challenged. Larry Temkin believes that the plausibility of the Levelling Down Objection depends on that of the *Person-Affecting Principle*, the principle that one outcome can be worse than another in some respect only if it is worse *for someone* in some respect. Temkin believes there are counterexamples to this principle that are independent of Egalitarianism.³ Parfit, however, holds that the Levelling Down Objection is more plausible than the Person-Affecting Principle and can be maintained without embracing the Person-Affecting Principle. That is, regardless of what one says about other alleged cases of non-person-affecting values, it is implausible in particular that an *inequality* is bad in cases in which that inequality is to the detriment of no one, not even of the least advantaged.⁴

This may serve as an overview of the state of play in regard to Egalitarianism and the Priority View. In the following, I suggest that both views are mistaken: the utility redistribution described in the opening paragraph above would not be an improvement. Egalitarianism and the Priority View are both exposed to an objection

² Derek Parfit, 'Equality or Priority?', *The Ideal of Equality*, ed. Matthew Clayton and Andrew Williams (New York, 2000), pp. 81–125; Thomas Nagel, 'Equality', *The Ideal of Equality*, pp. 60–80; Robert Young, 'Egalitarianism and Envy', *Philosophical Studies* 52 (1987), pp. 261–76; John Rawls, *A Theory of Justice*, rev. edn. (Cambridge, Mass., 1999), p. 72. Even Temkin recognizes the Priority View's intuitive plausibility ('Equality, Priority, and the Leveling Down Objection', *The Ideal of Equality*, pp. 126–61, at 130).

³ Temkin, 'Equality', pp. 137–40.

⁴ Parfit, 'Equality or Priority', p. 115; *Reasons and Persons* (Oxford, 1987), pp. 422–5.

that I shall refer to as the Pareto Argument. This argument is more compelling than the Levelling Down Objection and should be accepted by all who accept the Levelling Down Objection, in addition to many who do not accept the Levelling Down Objection.

II. AXIOLOGICAL REASONING: ASSUMPTIONS AND METHODS

My concern in the following is axiological, rather than deontic: I focus on what makes a possible state of affairs better or worse than another. This bears on questions of what we ought to do, since we generally have reason for promoting good states of affairs; however, no axiological claim by itself *entails* a conclusion as to what any agent ought to do, since there may be non-consequentialist reasons for action in addition to consequentialist reasons. I set aside any such non-consequentialist reasons; thus, I will not address such questions as whether we have duties of justice, or other non-consequentialist duties, either to try to bring about greater equality or to refrain from doing so. My central question is whether a state of affairs in which a given group of people – especially, the members of some society – are equally well off, is *pro tanto*, intrinsically better than a state of affairs in which the same group has a slightly higher total utility but with great inequality of welfare among the members of the group.

How should one reason about axiological questions of this kind? Three methodological cautions are in order. First, when reasoning about questions of intrinsic value, we must take care to avoid what Shelly Kagan calls ‘the additive fallacy’.⁵ One commits this fallacy when one argues that a certain factor is *in general* morally irrelevant, on the basis of an examination of a particular case, or a narrow class of cases, in which that factor fails to make a moral difference. This form of argument may lead us astray, for some factors may be morally relevant given certain background conditions, but irrelevant given other background conditions.

For example, consider a case in which I have promised to return a certain book to you. Then I learn that you stole the book from the library. In that case, some would argue, I should return the book to the library, and I have *no* reason, despite my promise, to give the book back to you.⁶ For the sake of argument, suppose that judgement is correct. In this case, my promise is normatively irrelevant. But it would be wrong to infer that promises *in general* are normatively irrelevant. In normal conditions, promises generate obligations, but in the particular

⁵ Shelly Kagan, ‘The Additive Fallacy’, *Ethics* 99 (1988), pp. 5–31.

⁶ Jonathan Dancy, *Moral Reasons* (Oxford, 1993), p. 60.

case in question, there are unusual, morally significant circumstances that prevent the promise from generating an obligation.

Keeping this in mind, when reasoning about the intrinsic value of equality, we would be wise to avoid relying solely on the consideration of cases that might be thought to exhibit unusual, morally significant features. Our arguments will be more cogent if we can rest them on the consideration of *paradigm* cases, that is, the typical sort of cases in which equality is thought to have value. If we argue that equality fails to enhance the value of a world in some abnormal case, that will not prevent the defender of equality from holding that equality *normally* contributes to the intrinsic value of the world. But if we argue that equality fails to enhance the value of a world in some normal case, a case of exactly the sort that Egalitarians have in mind when they speak of the intrinsic value of equality, this would pose a much more serious challenge to Egalitarianism. Similar points apply to reasoning concerning the Priority View.

So much for the first methodological principle. A second methodological principle concerns what sort of assumptions are admissible in philosophical reasoning. Sometimes an argument is open to criticism because one of its premises either depends for its justification on the conclusion, or simply is so similar to the conclusion that it is hard to imagine a reasonably clear-headed individual doubting the conclusion yet embracing the premise. In addition, sometimes arguments are subjected to criticism on the basis of a more demanding standard: sometimes it is suggested that an argument lacks merit because a clear-sighted, consistent opponent would not accept all of the argument's premises.

This last standard is unfairly demanding. To see why, let T be any theory, and let P be the conjunction of the premises of an argument against T. Suppose the argument is valid, so P entails $\sim T$. It follows that T entails $\sim P$. Given this, a committed proponent of T, who saw the logical consequences of his beliefs and was fully consistent, would not accept P, and thus would not accept all of the argument's premises. If this marks a flaw in the argument, then all valid arguments are *ipso facto* flawed. But a dialectical standard that rules out all valid arguments is unfairly demanding. Therefore, the requirement that the premises of an argument be acceptable to a logically clear-sighted, consistent opponent is an unfairly demanding dialectical standard.

That unfairly demanding standard results from a particular interpretation of the fallacy of begging the question, an interpretation that we can now see to be flawed. What interpretation might we put in its place? The rough intuitive idea behind strictures against question-begging is that someone presenting an argument for a conclusion may not assume that very conclusion. By the same token, one *responding*

to an argument may not assume the negation of the argument's conclusion. Thus, a fair constraint is that an argument must use premises that would seem plausible to most of those who consider those premises independent from the argument's conclusion – that is, to those who think about the premises without assuming either the truth of the conclusion or the falsity of the conclusion. An argument whose premises seem plausible only when one assumes the truth of the conclusion begs the question. An argument whose premises seem false when one assumes the falsity of the conclusion does not thereby beg the question; on the contrary, this is to be expected of a valid argument.

Finally, it is worth mentioning a methodological principle concerning the resolution of conflicts among intuitively plausible ethical principles. Suppose A and B each seem plausible independently, but they are found to be mutually inconsistent. One must be rejected. A number of criteria for choosing which principle to reject seem appropriate. Other things being equal, one should prefer to reject principles that seem less plausible, or less obvious, rather than those that seem more obvious; one should prefer to reject less widely shared intuitions, rather than those that are more widely shared; and one should prefer to reject intuitions that are more isolated from the rest of one's belief system (in the sense that their revision would require a smaller change in one's overall belief system), rather than those that are more tightly connected to one's belief system.

A special case of the second criterion occurs when evaluative intuitions vary significantly between different groups of people, where these groups can be identified independently of their intuitions on the particular matter in question. Suppose, for instance, that women tend to have the intuition that P, whereas men tend to have the intuition that \sim P. In that case, both P and \sim P would be shaky premises on which to rest philosophical conclusions. Whichever proposition one finds oneself inclined towards initially, one should be receptive to giving up one's intuition in favour of other intuitions that do not vary according to one's sex. Similarly, suppose that liberals tend to have the intuition that P, while conservatives tend to have the intuition that \sim P (where 'liberals' and 'conservatives' can each be identified by characteristics independent of attitudes toward P). In that case, again, both intuitions should be held open to revision, particularly on the grounds of other intuitions that do not vary according to political ideology.

To summarize our methodological principles: the best way to reason about whether some feature of a state of affairs contributes to intrinsic value is to (i) consider the central cases in which that feature is thought to matter, (ii) reason on the basis of premises that seem plausible when considered independent of the assumption that the feature in question does (or does not) contribute to intrinsic value, and (iii) avoid reliance

on controversial intuitions, in favour of intuitions whose appeal is more nearly universal.

III. THE PARETO ARGUMENT

III.1. A^+ is better than A

I begin by comparing two possible worlds, A and A^+ (figure 1). A contains one million people, each at the very high welfare level of 101. A^+ contains the same one million people, now at a welfare level of 102, plus an additional one million people at a welfare level of 1, just above the level at which life would cease to be worth living. Assume there are no further evaluatively relevant differences between A and A^+ not entailed by these descriptions. Call the well-off group that exists in both worlds ‘the Advantaged’, and call the worse-off group that exists in A^+ ‘the Disadvantaged’.

Which world is better, A or A^+ ? It seems to me that A^+ is better. A rough, intuitive motivation for this claim is that A^+ seems to be better *for everyone* than A is. It is better for the Advantaged, since A^+ gives them 102 units of utility, whereas A only gives them 101. It is also better for the Disadvantaged, since it is better to live at a welfare level of 1 than not to exist at all.

This reasoning is related to the reasoning of the Levelling Down Objection. The Levelling Down Objection assumes that, given two states of affairs x and y , if x is worse than y *for the Advantaged*, and x is not better for anyone in any respect, then x is not better than y *in any respect*. I find this principle plausible; however, my argument does not depend on it. My assumption is weaker in two important respects, each of which makes my argument harder to resist than the Levelling Down Objection.

First, those who advance the Levelling Down Objection find it implausible that, due to the value of equality, it can be desirable (in one respect) simply to harm the Advantaged. Be that as it may, it is surely even more implausible to hold that, due to the value of equality, it

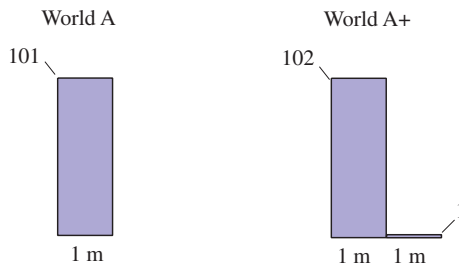


Figure 1.

can be desirable to harm *the Disadvantaged*, as well as the Advantaged. Admittedly, this is not precisely what is at issue, since the question here does not concern a *transition* from A^+ to A or vice versa, but rather a static comparison between two ways the world might have been. A more precise statement, then, is this: given two states of affairs x and y , if x is worse than y for *the Disadvantaged as well as for the Advantaged*, and x is not better than y for anyone, then x is not better than y .

The second way in which my assumption is weaker than that involved in the Levelling Down Objection is that the Levelling Down Objection relies on an intuition that a certain kind of situation is not better than another *in any respect*. An Egalitarian might maintain that although a change that harms the better-off and helps no one is good *in one respect*, this is of course outweighed by the obvious respect in which the change is bad, leaving it bad overall. To some degree, this mitigates the implausibility of the consequence drawn out by the proponents of the Levelling Down Objection. The Egalitarian could not, however, take an analogous fall-back position in responding to my argument here. For my reasoning does not rest on the strong claim that A is not better than A^+ *in any respect*. Perhaps there are respects in which A is better than A^+ . I deny only that A is better than A^+ *overall*.

Now consider an objection to the preceding reasoning. I have suggested that A^+ is better than A because it is better both for the Advantaged and for the Disadvantaged. But some would argue that A^+ is not better for the Disadvantaged, for the following reason. For A^+ to be better than A for the Disadvantaged, the Disadvantaged must have a higher welfare level in A^+ than they have in A . Since the Disadvantaged do not exist in world A , they have no welfare level at all (not even level 0) in world A . Therefore, it is not the case that they have a higher welfare level in A^+ than they have in A . Rather than concluding that A^+ is better for the Disadvantaged than A is, we should conclude that A and A^+ are incomparable from the standpoint of the interests of the Disadvantaged.⁷

There are three promising ways of responding to this objection. The first is to point out that, even if A^+ is not better than A for the Disadvantaged, it is certainly no worse for them. Since A^+ is better for some (the Advantaged) and worse for no one, it remains plausible that A^+ is better than A overall.

The second way of responding is to argue that a state of affairs in which one does not exist is, in fact, one in which one has zero utility.

⁷ Jan Narveson, 'Utilitarianism and New Generations', *Mind* 76 (1967), pp. 62–72, at 67; Jan Narveson 'Moral Problems of Population', *The Monist* 57 (1973), pp. 62–86, at 66; see also John Broome, 'The Welfare Economics of Population', *Oxford Economic Papers* 48 (1996), pp. 177–93, at 179.

Although it is obvious that one cannot have any non-zero amount of utility when one does not exist, it is not obvious that one cannot have zero utility when one does not exist. One might think that not existing entails not having any utility or disutility, and that not having any utility or disutility entails having zero utility.

Third, and most importantly, even if it is not literally true that one has zero utility when one does not exist, it is certainly true that one does not have either utility or disutility when one does not exist, and it is plausible to regard a state in which one has neither utility nor disutility as evaluatively equivalent to a state in which one has welfare level zero. Three considerations support this contention:

- a. Intuitively, it is *worse* to exist at a *negative* welfare level (for instance, living a life of constant agony) than not to exist at all, and nearly everyone would judge a world containing some people with horrible lives to be *pro tanto* worse than one in which those people do not exist at all. Symmetry considerations thus suggest that it is *better* to exist at a *positive* welfare level than not to exist at all, and that we should judge a world containing some people with valuable lives to be *pro tanto* better than one in which those people do not exist at all.
- b. It is prudentially rational to prefer a *future* in which one continues to exist at a positive welfare level, over one in which one ceases to exist, over one in which one continues to exist at a negative welfare level. This suggests that, from the standpoint of self-interest, it is rational to treat a state of affairs in which one does not exist as equivalent to one in which one has zero welfare.
- c. Imagine asking a fully informed, rational, purely self-interested individual whether he was glad that he existed, or whether he wished he had not been born. Provided this individual had (and expected to have over his lifetime) a positive welfare level, he would presumably answer that he was glad he existed; if he had (and expected to have over his lifetime) a negative welfare level, he would presumably answer that he wished he had not been born. These attitudes are surely intelligible, whatever one may think about the question of whether one can be *better off* in a situation in which one exists than in a situation in which one does not exist. But reflection on these attitudes suggests, again, that for evaluative purposes, we treat non-existence as equivalent to a state of zero welfare.⁸

⁸ I defend this point at greater length in my 'In Defence of Repugnance', *Mind* 117 (2008), pp. 899–933.

In light of the immediately preceding considerations, we may wish to replace talk about what is *better for* the Disadvantaged with talk about what the Disadvantaged, if fully informed, rational and self-interested, *would prefer*. Thus, the rationale for judging world A^+ better than world A can be expressed by the following principle:

The Weak Pareto Principle: For any two possible worlds, x and y , if all of the following conditions hold –

- (i) x has greater total utility than y ,
- (ii) x would be preferred to y , from the standpoint of fully informed, rational self-interest, by every individual who would exist in either world,
- (iii) all the inhabitants of x deserve the benefits they receive in x , or at least do not deserve *not* to receive them, and
- (iv) there are no morally relevant differences between x and y apart from differences in their utility distributions (for example, they do not differ in amounts of virtue, or knowledge, or freedom)⁹

– then x is better than y .

Clause (iii) is needed to rule out the possible situation in which some inhabitants of world x deserve to suffer (perhaps for retributivist reasons), in which case their being better off might be a bad-making feature of x . The Weak Pareto Principle reflects the intuitive idea that (provided we are dealing with normal, not undeserving people) our evaluations of different possible distributions of utility should be guided by a spirit of benevolence: what is good for all must be good.

If we agree, then, that A^+ is better than A , we might conclude from this that Egalitarianism is false. For A exhibits perfect equality, while A^+ exhibits gross inequality, and A^+ has only slightly greater total utility. If, despite this, A^+ is better than A , then equality must have no significant value.¹⁰

But in view of the first methodological stricture of section II above, the preceding argument appears too hasty. Questions of population ethics, involving what Parfit calls ‘different number choices’, are notoriously puzzling.¹¹ An Egalitarian might well feel that little weight should be placed on cases in which we compare worlds containing different numbers of people. Perhaps in some such cases, inequality does not

⁹ A ‘utility distribution’ is a specification of how many individuals exist at each welfare level in a given possible world or state of affairs.

¹⁰ Since qualitatively similar examples can be constructed in which A^+ has arbitrarily small advantages in total utility over A , and yet A^+ remains better than A , we can infer that the value of equality must be zero.

¹¹ See Parfit, *Reasons and Persons*, p. 356, ch. 17.

matter, but in other, more standard cases, inequality is still intrinsically bad.¹² The standard sort of case envisioned by those who discuss the value of equality is one in which a fixed group of individuals may have either equal or unequal utility levels. Most of those with Egalitarian intuitions feel – when considering, for instance, the welfare gap between the rich and the poor – that things would be better if only some of the benefits enjoyed by the wealthy could be given to the poor. But they do not typically feel that things would be better if only the individuals who are poor *did not exist*. Most Egalitarians would like to eliminate poverty through giving more wealth to the poor; they would not like to eliminate poverty through eliminating the poor. Thus, it is plausible that an Egalitarian would deny that Egalitarianism implies that world A is better than A⁺. Instead, the Egalitarian might reasonably maintain that inequality matters when comparing possible distributions of utility across a fixed group of people, but that it does not necessarily matter when comparing different possible populations of people.

For this reason, we have not yet found a compelling argument against Egalitarianism. Unless we can show that inequality does not matter, even when considering utility distributions across a fixed group of individuals, we will not have refuted the core contention of Egalitarianism.

Even more clearly, the contention that A⁺ is better than A is perfectly compatible with the Priority View. Those who embrace the Priority View hold that benefiting anyone is good, but benefiting those who are presently disadvantaged is more important than benefiting those who are presently advantaged. This view, if anything, suggests an acceptance of the Weak Pareto Principle, which entails that A⁺ is better than A.

III.2. *A is better than B*

Consider another pair of worlds, A and B (figure 2). World A is as described before. World B contains the same two million people as in A⁺, but each at a welfare level of 50. There are no other evaluatively relevant differences between A and B. Which world is better?

This is a question of population axiology. It is therefore appropriate to review the prominent, recently discussed theories of population axiology, and see what each theory implies about our question. There

¹² For example, Andrew Mason proposes that inequality is intrinsically bad, but only in circumstances in which it is to the disadvantage of someone ('Egalitarianism and the Levelling Down Objection', *Analysis* 61 (2001), pp. 246–54, at 248).

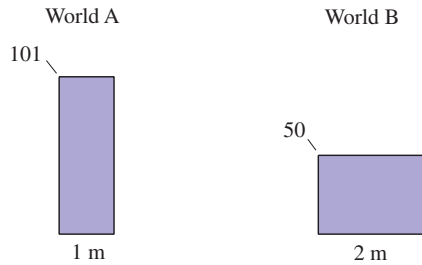


Figure 2.

are eight such theories:

1. *Average Utilitarianism* holds that the better world is the one whose members have the higher average level of well-being. This is world A.
2. *Total Utilitarianism* holds that the better world is the one whose total utility is higher. This is again world A.
3. *Critical Level Theories* hold that there is some threshold level of welfare (above zero) at which a life begins to add to the value of the world – if one's welfare is between zero and the critical level, one's life has value to oneself but does not make the world as a whole better.¹³ This view would make a difference to the evaluation of worlds A and B only in the (unlikely) event that the critical level were at or above 50, in which case the superiority of A would only be made all the more clear, since none of the lives in B would contribute anything to the world's value.
4. *Perfectionism* holds that 'even if some change brings a great net benefit to those who are affected, it is a change for the worse if it involves the loss of one of the best things in life'.¹⁴ This claim matters to the comparison of A and B only if world A contains one of the best things in life while B does not (we can stipulate that the reverse is not the case – it is not the case that B contains one of the best things in life while A does not); in this case, Perfectionism would only reinforce the judgement that A is better than B.
5. *The Person-Affecting Principle* holds that one world can be better than another only if it is better for someone.¹⁵ World A is better

¹³ Charles Blackorby, Walter Bossert and David Donaldson, 'Critical-Level Population Principles and the Repugnant Conclusion', *The Repugnant Conclusion: Essays on Population Ethics*, ed. Jesper Ryberg and Torbjorn Tännsjö (Dordrecht, 2004), pp. 45–59.

¹⁴ Derek Parfit, 'Overpopulation and the Quality of Life', *The Repugnant Conclusion: Essays on Population Ethics*, pp. 7–22, at 19.

¹⁵ Narveson, 'Utilitarianism'; Narveson, 'Moral Problems'; Broome, 'Welfare Economics'. For further discussion of the Person-Affecting Principle, see Parfit, *Reasons*

for the Advantaged than B is, so the Person-Affecting Principle poses no objection to regarding A as superior to B.

6. *Variable Value Theories* hold that the value that an individual's life contributes to a world depends not just on that individual's utility, but also on the number of other people who exist. The theories of this kind that have actually been advanced all depict the value of additional lives diminishing with population size – that is, the more people there already are, the less improvement is made to the world by the addition of a new person at a given welfare level.¹⁶ Such theories only reinforce the judgement that A is better than B.
7. *Lexicality* holds that a fixed duration of very intense pleasure is superior to any amount of very mild pleasure, no matter how protracted the mild pleasure may be or how many individuals enjoy it.¹⁷ This view, again, could only reinforce the preference for A over B.
8. *Justicism* holds that for each person there is a level of welfare that person deserves. The value of a given person's enjoying some welfare level is diminished according to how far that welfare level is from the level he deserves. The value of a world is the sum of the values of the episodes of well-being in that world.¹⁸ To assess whether A is better than B on this view, we must stipulate a level of desert for the inhabitants of each world – let us stipulate that, in each of worlds A, B and A⁺, all persons deserve 102 units of well-being. In this case, B is clearly worse than A, since B's total utility is lower and its inhabitants are much farther from their deserved welfare level.¹⁹

Most of the above theories are designed specifically to avoid the Repugnant Conclusion, namely, the conclusion that for any given world

and Persons, p. 394, and Larry Temkin, 'Intransitivity and the Mere Addition Paradox', *Philosophy and Public Affairs* 16 (1987), pp. 138–87, at 166–7.

¹⁶ Yew-Kwang Ng, 'What Should We Do About Future Generations? Impossibility of Parfit's Theory X', *Economics and Philosophy* 5 (1989), pp. 135–253; Thomas Hurka, 'Value and Population Size', *Ethics* 93 (1983), pp. 496–507; Ted Sider, 'Might Theory X Be a Theory of Diminishing Marginal Value?', *Analysis* 51 (1991), pp. 265–71.

¹⁷ Stuart Rachels, 'Counterexamples to the Transitivity of *Better Than*', *Australasian Journal of Philosophy* 76 (1998), pp. 71–83; 'A Set of Solutions to Parfit's Problems', *Nous* 35 (2001), pp. 214–38; Larry Temkin, 'A Continuum Argument for Intransitivity', *Philosophy and Public Affairs* 25 (1996), pp. 175–210.

¹⁸ Fred Feldman, 'Justice, Desert, and the Repugnant Conclusion', *Utilitarianism, Hedonism, and Desert* (Cambridge, 1997), pp. 195–214.

¹⁹ This stipulation also has the result that, according to justicism, world A⁺ would be worse than A (cf. Feldman, 'Justice', pp. 209–10). However, I remain persuaded, because of the considerations of section III.1 above, that justicism is wrong about this. Someone's being extremely praiseworthy and hence highly deserving cannot make it the case that it would be better if he did not exist, rather than existing and receiving a small benefit.

full of happy people, a world containing a sufficiently large number of people with lives barely worth living would be better.²⁰ It is not surprising that theories designed to avoid the Repugnant Conclusion dictate a preference for A over B. But some philosophers embrace the Repugnant Conclusion.²¹ They do not do this because they value population size intrinsically. Rather, these philosophers explain the truth of the Repugnant Conclusion in terms of the value of (total) utility. The most extreme position anyone takes in this direction – the position favouring larger populations in the greatest range of cases – is the Total Utility Principle. But world B is inferior to A even in respect of total utility. Thus, every contemporary theory in population axiology converges on the judgement that world A is better than world B.

The principle behind this step in the argument can be formulated as follows:

The Unrepugnant Premise: For any two possible worlds, x and y , if all of the following conditions hold –

- (i) x has greater total utility than y ,
 - (ii) x has higher average utility than y ,
 - (iii) neither x nor y has any inequality, and
 - (iv) there are no relevant differences between x and y apart from their utility distributions
- then x is better than y .

III.3. A^+ is better than B

We have said that A^+ is better than A, which is better than B. Therefore, A^+ is better than B (figure 3).

This conclusion is at odds with both Egalitarianism and the Priority View. Worlds A^+ and B contain the same individuals. A^+ has only slightly higher average and total utility than B, but A^+ exhibits gross inequality, while B exhibits perfect equality. On Egalitarianism, this

²⁰ I have formulated the Repugnant Conclusion slightly differently from Parfit's (*Reasons and Persons*, p. 388) original formulation.

²¹ Bill Anglin, 'The Repugnant Conclusion', *Canadian Journal of Philosophy* 7 (1977), pp. 745–54; R. I. Sikora, 'Is It Wrong to Prevent the Existence of Future Generations?', *Obligations to Future Generations*, ed. R. I. Sikora and Brian Barry (Philadelphia, 1978), pp. 112–66; Ng, 'Future Generations'; Ng, 'Welfarism and Utilitarianism: A Rehabilitation', *Utilitas* 2 (1990), pp. 171–93; Robin Attfield, *The Ethics of Environmental Concern*, 2nd edn. (Athens, Ga., 1991), pp. 127–30; Jesper Ryberg, 'Is the Repugnant Conclusion Repugnant?', *Philosophical Papers* 25 (1996), pp. 161–77; Nick Fotion, 'Repugnant Thoughts About the Repugnant Conclusion Argument', *Contingent Future Persons*, ed. Nick Fotion and Jan C. Heller (Dordrecht, 1997), pp. 85–97; Torbjörn Tännsjö, 'Why We Ought to Accept the Repugnant Conclusion', *Utilitas* 14 (2002), pp. 339–59; John Broome, *Weighing Lives* (Oxford, 2004), pp. 210–14; Huemer, 'In Defence of Repugnance'.

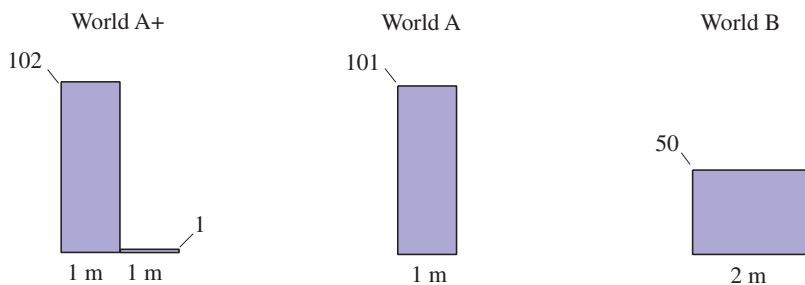


Figure 3.

should make B better than A⁺. On the Priority View, the Disadvantaged in A⁺ should take priority over the Advantaged, so that raising the utility of the Disadvantaged by a given amount would be better than avoiding the loss of a slightly greater amount of utility by the Advantaged. Thus, a change from A⁺ to B should be a change for the better. If, as I have argued, A⁺ is better than B, then both Egalitarianism and the Priority View are mistaken.

The inference to the conclusion that A⁺ is better than B depends on the assumption that the better-than relation is transitive:

Transitivity: For any x , y , and z , if x is better than y and y is better than z , then x is better than z .

This assumption enjoys strong, widespread intuitive appeal. Nevertheless, some have rejected it.²² Larry Temkin in particular argues against Transitivity on the basis of Egalitarianism.²³ Because of this, it is worth adducing arguments for Transitivity, beyond the direct appeal to intuition.

The best-known argument for Transitivity is the Money Pump Argument.²⁴ Suppose that X is better than Y , which is better than Z , which is better than X . It seems that a rational person might then prefer X to Y , Y to Z , and Z to X . Suppose you are such a person. You presently have X . I offer to let you pay a small amount of money to be allowed to trade X for Z . Since you prefer Z , you accept. I then let you pay a small amount of money to trade Z for Y . Again you accept. I then let you pay a small amount of money to trade Y for X . You accept. You thus wind up back where you started, only with less money. Since you

²² See Rachels, 'Counterexamples', and Temkin, 'Continuum' (giving a variant of Rachels' argument).

²³ Temkin, 'Intransitivity', pp. 171–3.

²⁴ This argument is from Donald Davidson, J. C. C. McKinsey and Patrick Suppes ('Outlines of a Formal Theory of Value', *Philosophy of Science* 22 (1955), pp. 140–60), who credit Norman Dalkey.

predictably reach this situation solely by acting on your preferences, it seems that those preferences are not rational, contrary to our initial supposition.²⁵

A second argument for Transitivity relies on the following two premises:

Dominance: Given two sets of states of affairs, $\{x_1, \dots, x_n\}$ and $\{y_1, \dots, y_n\}$, if

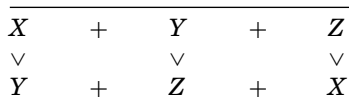
- (i) x_i is better than y_i for every i , and
- (ii) either there are no evaluatively significant relationships among any of these states, or any evaluatively significant relationships among the x 's are mirrored among the y 's,

then the combination of x_1, \dots , and x_n is better than the combination of y_1, \dots , and y_n .

Irreflexivity: Nothing is better than itself.

To illustrate the Dominance principle, suppose that I am deciding whether to buy a Honda or a Ford. I am also deciding whether to live in California or Texas. Assume there are no evaluatively significant relationships between these choices: where I live has nothing to do with what kind of car is best, and vice versa. Finally, suppose that the Honda is better than the Ford, and living in California is better than living in Texas. Then it seems that buying the Honda and living in California would be better than buying the Ford and living in Texas.

Now suppose that Transitivity is false, and that there is a set of unrelated states, X , Y and Z , such that X is better than Y , which is better than Z , which is better than X . I shall denote the combination of X , Y and Z , ' $X + Y + Z$ ' – this is the state of affairs that obtains precisely when X , Y and Z all obtain. Now consider which is better: $X + Y + Z$, or $Y + Z + X$? By Dominance, $X + Y + Z$ is better than $Y + Z + X$, because X is better than Y , Y is better than Z , and Z is better than X . The following diagram depicts these relationships:



²⁵ Robert Nozick, (*The Nature of Rationality* [Princeton, NJ, 1993], p. 140 n.), Stuart Rachels ('Counterexamples', pp. 82–3), and Chrisoula Andreou ('Environmental Damage and the Paradox of the Self-Torturer', *Philosophy and Public Affairs* 34 (2006), pp. 95–108) question the implicit premise that it is always rational to choose the better of two options when one knows which option is better. None, however, appear to offer grounds for doubting this premise that are independent of the assumption of Intransitivity.

But this is absurd. $X + Y + Z$ is the same state as $Y + Z + X$; this state cannot be better than itself. The most plausible candidate for what went wrong in the foregoing reasoning is the assumption that X , Y and Z might have an intransitive evaluative ordering.

IV. ASSESSING THE PARETO ARGUMENT

The Pareto Argument, in sum, goes as follows:

1. A^+ is better than A (by the Weak Pareto Principle).
2. A is better than B (by the Unrepugnant Premise).
3. If A^+ is better than A and A is better than B , then A^+ is better than B (by Transitivity).
4. Therefore, A^+ is better than B (from 1, 2, 3).

This argument is obviously deductively valid. It also avoids the Additive Fallacy, discussed in section II. The comparison between worlds A^+ and B is a paradigm of the kind of case in which Egalitarians and Prioritarians hold a more equal distribution to be intrinsically better. If equality does not matter to the comparison of *these* two worlds, then it is safe to conclude that equality never matters.

We have discussed above the plausibility of each of the argument's premises. But are there ways in which an Egalitarian or Prioritarian might object to one or more premises? Begin with premise (1), that A^+ is better than A . An Egalitarian might argue that A is better than A^+ , because A exhibits perfect equality, A^+ contains gross inequality, and A^+ has only trivially greater total utility. Because Egalitarianism may lead to the denial of premise (1) in such a simple and obvious way, it might be thought that premise (1) begs the question against Egalitarians.

If premise (1) had been asserted without argument, this objection might be fair. Premise (1), however, was supported by further reasoning, involving the Weak Pareto Principle. The Weak Pareto Principle is intuitively highly plausible. To deny it in this case, the Egalitarian would have to hold that the value of equality can sometimes justify preferring a state of affairs that is worse even for the Disadvantaged than another (or at least would be dispreferred by the Disadvantaged). Even those with initial Egalitarian intuitions may balk at this.

Nevertheless, suppose that the Egalitarian rejects the Weak Pareto Principle, on the grounds that it conflicts with Egalitarianism in such cases as that of the comparison between worlds A and A^+ . Would this constitute a convincing response to the Pareto Argument against Egalitarianism?

Recall the methodological strictures advanced in section II. If premise (1) or the Weak Pareto Principle had been justified by appeal to the denial of Egalitarianism, then the Pareto Argument would certainly beg the question. This, however, is not the case; the Weak Pareto Principle has independent, intuitive appeal. When we consider the Weak Pareto Principle apart from Egalitarianism (without assuming Egalitarianism or its negation), the principle seems plausible. Its use in an argument against Egalitarianism therefore does not constitute an objectionable begging of the question. Rather, the *objection* to the Weak Pareto Principle entertained here would beg the question. We have imagined the Egalitarian objecting to the Weak Pareto Principle solely on the basis of Egalitarianism itself. Since the Pareto Argument is an argument against Egalitarianism, this is a dialectically inappropriate objection: the Pareto Argument cannot fairly be evaluated by first assuming the negation of its conclusion. To fairly object to the Weak Pareto Principle, the Egalitarian must adduce grounds for doubting the Weak Pareto Principle that do not depend on the truth of Egalitarianism.

Consider next premise (2), that A is better than B. In this case, a Prioritarian might object on the following grounds. The most straightforward way of cashing out the Priority View as an axiological thesis is to hold that utility has diminishing marginal value: that is, the higher an individual's welfare level is, the less agent-neutral value is produced by a given, fixed-size addition to that individual's welfare.²⁶ This formulation of the Priority View leads directly to the conclusion that world B is better than A (provided the rate at which the value of utility diminishes is non-trivial), for the value of a life at welfare level 101 will be less than twice that of a life at welfare level 50. Since world B has twice as many lives (at welfare level 50) as world A has (at welfare level 101), world B will have more value, assuming that value is additive across persons. Since the Priority View leads to the denial of premise (2) in this fairly straightforward way, it might be argued that premise (2) begs the question in an argument against the Priority View.

Again, it is the objection to the Pareto Argument, rather than the Pareto Argument itself, that begs the question. Premise (2) was not justified by an appeal to the negation of the Priority View. Instead, premise (2) accords with widely shared intuitions in population axiology, as well as all the major current theories of population axiology (with the exception of the Priority View itself, if we consider that as a theory of population axiology). Many of these theories have been

²⁶ This is explicit in John Broome's formulation of the Priority View (*Weighing Goods: Equality, Uncertainty and Time* (Oxford, 1991), pp. 179, 199–200).

devised by philosophers seeking to avoid the Repugnant Conclusion; none is motivated by a desire to avoid the Priority View. One of the theories (the Total Utility Principle) accepts the Repugnant Conclusion; but this theory is not motivated by a desire to avoid the Priority View either. It is typically motivated by a desire to avoid other paradoxical consequences that seem even less plausible than the Repugnant Conclusion.²⁷ Thus, the appeal to premise (2) does not beg the question against the Priority View. If the only objection to premise (2) stems from the Priority View itself, then this objection begs the question. The Pareto Argument is an argument against the Priority View (as well as Egalitarianism); it cannot be fairly evaluated by reasoning from the assumption that the Priority View is correct.

The following five principles each have some intuitive plausibility:

- The Weak Pareto Principle
- The Unrepugnant Premise
- Transitivity
- Egalitarianism
- The Priority View

But all five cannot be correct. The first three conjointly conflict with each of the last two. We must therefore reject one of the first three principles, or else reject both of the last two.

The first three principles enjoy widespread intuitive appeal. These intuitions seem to derive from rational reflection and seem unaffected by, for example, one's sex, social class, or political ideology. Admittedly, those who embrace Egalitarianism or the Priority View may, after learning of the Pareto Argument, be inclined to reject one of the first three principles. But even those with Egalitarian or Prioritarian sympathies will most likely feel the force of the intuitions, and would most likely embrace all of the first three principles, provided that they did not see how the principles conflict with Egalitarianism and the Priority View.

Similar remarks cannot be made about the last two principles. Egalitarianism and the Priority View each have a narrower intuitive appeal – independent of the argument of this article, both are controversial. In particular, partisans of opposing political ideologies tend to disagree widely in their assessments of the plausibility of views in the neighbourhood of Egalitarianism (including the Priority View). Those on the political left tend to find such views much more initially plausible than do those on the political right. Some see in the appeal of Egalitarianism nothing better than envy, as indicated in Oliver Wendell

²⁷ See the sources cited in n. 21 above.

Holmes's remark: 'I have no respect for the passion for equality, which seems to me merely idealizing envy'.²⁸

The point here is not that Holmes is correct in his assessment of the source of Egalitarian intuitions. Holmes's assessment remains speculative and open to debate. Rather, the point is that, prior to the consideration of the argument of this article, the plausibility of Egalitarianism and related views is in dispute. But there is no such dispute about the plausibility of the Weak Pareto Principle, the Unrepugnant Premise, or Transitivity. This does not foreclose the possibility of further ethical arguments against one or more of those three principles. But one ought not to reject such widely shared intuitions solely on the basis of other, more controversial intuitions.

Controversial intuitions – such as the intuition that welfare inequality is intrinsically bad or that benefiting the worse-off is more important than benefiting the better-off by the same amount – should be tested by how well they cohere with other, less controversial intuitions – such as that what is better for everyone is better overall, or that it is worse to have a larger population with lower average and total utility, or that *better-than* is transitive. The most that can generally be expected of ethical reasoning is that it show some controversial ethical thesis to conflict with less controversial intuitions. This is what the Pareto Argument accomplishes. This provides us with grounds for denying Egalitarianism and the Priority View. These grounds are not entirely conclusive, since ethical intuitions are fallible – but the grounds are as cogent as can fairly be expected. Unless some further, independent grounds for denying one of the argument's premises are forthcoming, we should reject both Egalitarianism and the Priority View.²⁹

owl232@earthlink.net

²⁸ Oliver Wendell Holmes and Harold J. Laski, *The Holmes–Laski Letters*, vol. 2, ed. Mark DeWolfe Howe (Cambridge, Mass., 1953), p. 942. Compare Helmut Schoeck, *Envy: A Theory of Social Behavior* (New York, 1970), p. 231: '[T]he sense of equity, of justice and injustice[, is] inherent in man because of his capacity to envy'; Friedrich Hayek, *The Constitution of Liberty* (Chicago, 1960), p. 93: '[M]ost of the strictly egalitarian demands are based on nothing better than envy'. In my own informal surveys, about a third of undergraduate students feel that inequality is intrinsically bad, another third are unsure, and the rest feel that inequality is not intrinsically bad. This result must be taken with a grain of salt due to the surveys' unscientific nature, as well as uncertainty regarding respondents' understanding of the question. Nevertheless, I am inclined to doubt that the Egalitarian intuition is extremely widespread.

²⁹ I would like to thank David Schmitz, the other philosophers at the University of Arizona (where this article was presented as a paper) and the participants of the PEA Soup weblog for some stimulating discussions of the argument of this article.