Saving Posterity from a Worse Fate

Niko Kolodny

August 22, 2020

1 Introduction

Suppose we must choose among different outcomes, in which people fare better or worse.[1] Suppose different numbers of people, or at least different people, would ever exist in such outcomes. That is, suppose our choice affects the growth of the population, or the identities of future people. Which outcomes are wrong for us to choose?

Consider the simplest sort of case. Suppose that Eleanor, but no one else, will exist whatever we choose. What's up to us is merely whether her neighborhood will be the Good Place, where she will flourish, or the Medium Place, where she will get by. It would be wrong, we might think, to put her in the Medium Place.

Some qualifications right away. First, the fact that someone will flourish or get by or suffer, fare well or so-so or badly, if we make a choice is a less compelling reason, if any reason at all, when they will fare badly only because of their own choices, despite our best efforts to provision them, to present them with good opportunities, to put them in a position to fare well, should they make the right choices. When I speak in this paper of our choice affecting how people fare, therefore, it should be understood as shorthand for affecting how people are provisioned, what opportunities they have access to. That, rather than how they ultimately fare, is what makes a claim on us.

Second, of course, we may have agent-relative reasons to give greater weight to ourselves, to our projects, and to our relationships. We may be prohibited from taking certain means to otherwise desirable outcomes. And there may be impersonal values — such as the progress of art and science, biodiversity and the health of natural ecosystems — that bear on our choice, apart from their effects on how people fare. But let's set these aside for the moment. Restricting our attention to how people fare in the outcomes that we might choose, which outcomes are wrong for us to choose?

Third, we might well think that it isn't wrong to put Eleanor in the Medium Place, if she wouldn't be too badly off there: if, to put it crudely, she would be above a threshold of sufficiency so that failing to raise her even higher, even at no cost to us, would be at most supererogatory. But suppose that she would be below that threshold in the Medium Place, although she's above the threshold of being glad to have been born.[2]

Granting that it is, intuitively, wrong to put Eleanor in the Medium Place, why is it? There are two broad ways of thinking about it. The more familiar way might be called "Benefit Thinking." Benefit Thinking explains why it is wrong to put Eleanor in the Medium Place by pointing out that we benefit people (in this case, Eleanor) more by putting her in the Good Place than we benefit people (again, still just Eleanor) by putting her in the Medium Place.

It's another way of thinking about it, which might be called, "Worse-Fate Thinking," that I want to explore. Worse-Fate Thinking explains why it would be wrong to put Eleanor in the Medium Place by pointing out that people in the Medium Place (namely, Eleanor) are left to a worse fate in the Medium Place. Put another way, the people in the Medium Place (namely, Eleanor) have reason to prefer how things are for people (not necessarily themselves, but in this case, as it happens, herself, Eleanor) in some alternative outcome that we might choose: namely, the Good Place. Put still another way, leaving Eleanor in the Medium Place is wrong because increases, relative to some alternative that we might choose, the number of people fated to exist a lower level of well-being. If we choose the Good Place, the number of people at every level less than whatever level of flourishing obtains at the Good Place is zero. If we choose the Medium Place, by contrast, the number at some level less than that level of flourishing increases to one.

One's first thought may be that the second answer, that we are saving people from something worse, has greater moral force. Isn't it more urgent to save people from something worse than merely to benefit them? But one's second thought may be to discount the first thought, as an illusion or artifact of framing. Surely benefitting Eleanor by putting her in the Good Place rather than the Medium Place comes to the same thing as saving her from the worse fate of the Medium Place by putting her in the Good Place. So Worse-Fate Thinking is no different from Benefit Thinking.

It is true that Worse-Fate Thinking and Benefit Thinking deliver the same verdicts in the ordinary, same-people choices with which we are most familiar, where neither the number nor the identities of people are affected by what we choose. But Worse-Fate Thinking delivers different verdicts from Benefit Thinking in more exotic choices that

---

[2]It might also be said that, while the fact that people would fare better or worse can be a reason for action, it is in many contexts not a reason of individual morality, which might tend to make a private person's choice wrong, but instead a reason of administrative morality, which would tend instead to make the choice of some public agency wrong or in some other sense incorrect. Or it might be said that it is not a moral reason at all, but instead an "extramoral" reason. For powerful arguments for these alternative interpretations of the significance, in many contexts, of the fact that people would fare better or worse, see (Wallace 2019). It may be that this paper is best understood as an account of administrative morality or extramoral reasons.

affect the numbers or identities of people.

The key difference between Worse-Fate and Benefit Thinking might be put this way. Suppose our choice is whether to cause Eleanor to exist at the Good Place or not to cause anyone to exist. In this case, choosing that no one exist does not save people from a worse fate. This is because existing with a life worth living is not a worse fate than not existing. Someone who exists with a life worth living does not have reason to prefer how things are for people when no people exist. So Worse-Fate Thinking finds no grounds to count as wrong causing Eleanor to exist. Benefit Thinking agrees, insofar as causing someone to exist with a life worth living benefits them. However, Eleanor's existing also does not save from a worse fate. Not to exist is not to suffer a worse fate worse than a life worth living. It's not a condition someone is in at all and so not a condition such that in it someone has reason to prefer how things are for people in some alternative. So Worse-Fate Thinking also finds no grounds to count as wrong not causing Eleanor to exist. Here Benefit Thinking disagrees, since not causing Eleanor to exist passes up a chance to benefit someone.

I will argue that the principle, "Worse-Fate," that I formulate to represent this alternative to Benefit Thinking gives intuitive answers to many of the questions of population ethics. To indulge in some jargon that will be familiar to the initiated, but that will have be explained in what follows, Worse-Fate accounts for the "procreation asymmetry," avoids the "repugnant choice," and delivers plausible results in "mixed same-number" cases.

What's the catch? One catch, it might be said, is that Worse-Fate implies that a choice may not be required, and may even be wrong, when it benefits people more. But this is just to say that Worse-Fate Thinking diverges from Benefit Thinking. And it is surprisingly hard to come up with non-question-begging grounds for the conclusion that our concern for how people fare reflects Benefit Thinking rather than Worse-Fate Thinking. After all, if we restrict our attention to ordinary, same-people choices, they seem to have equal title to represent our concern for how people fare, since they deliver precisely the same verdicts. If, once we consider the more exotic choices, which do affect numbers and identities, Worse-Fate Thinking delivers more intuitive verdicts, then why not conclude that Worse-Fate Thinking best represents our concern for how people fare? Why not conclude that we misinterpret ourselves when we suppose that our intuitions in the original case of Eleanor express Benefit Thinking? Why not conclude that our intuitions are in fact guided by Worse-Fate Thinking? Or, to consider an alternative hypothesis, perhaps both Worse-Fate and Benefit Thinking are operative in our concern for how people fare in ordinary, same-people choices. We just don't notice the difference in ordinary, same-people choices, since they deliver the same practical verdicts. To represent our concern for how people fare, we must have a principle that does justice to both Worse-Fate and Benefit Thinking. In that case, Worse-Fate can be amended, into a principle I call "Worse-Fate-Pareto." That avoids the most persuasive counterexamples (or, as it seems to me, apparent counterexamples) to Worse-Fate.

A second catch is that Worse-Fate gives practical verdicts that pattern in seemingly "incoherent" ways, such as violating "transitivity." However, I don't see a compelling

reason to assume that practical verdicts must be "coherent" in the alleged ways. So long as we can state a principle that explains the intuitively correct verdicts and that has an attractive rationale, we secure the sort of coherence that matters.

Worse-Fate implies that a choice to cause someone to be can be wrong even though it leaves no particular person worse off than that person would have been in any alternative. Suppose that our choice is either to cause Mindy to be in the Medium Place or to cause Eleanor to be in the Good Place. According to Worse-Fate, it would be wrong to cause Mindy to be in the Medium Place. But Mindy is not worse off than she would be in the alternative. How then can she complain? And if no one can complain, how can it be wrong? This is Parfit (1984)'s "non-identity problem."

However, Worse-Fate can be seen as part of a more complex view, Worse-Fate-As-Answer, where Worse-Fate functions as an answer to a person's putative complaint about having been caused to be. On this view, people can complain about choices that cause them to be, even when those choices do not make them worse off, on the grounds that such choices harm or invade them, without their consent. One thing that can answer such a putative complaint is that the choice saved people from a worse fate. However, when a choice to cause someone to be failed to save people from a worse fate, the people caused to be have an unanswered complaint against that choice. So Mindy in the Medium Place has a complaint if she is caused to be that Eleanor in the Good Place lacks if she is caused to be.

Whether we can accept any of these ideas, Worse-Fate, Worse-Fate-Pareto, or Worse-Fate-As-Answer, I don't know. Perhaps we are committed, and rightly so, to Benefit Thinking. But if so, then this paper at very least sets up a collision with error that gives us a sharper appreciation of that deep truth.

2 Taxonomy

Some labels will help us to keep things straight. Say that someone is "contingent" relative to a choice set if whether that person ever exists depends on the choice we make: there's at least one choice in the set where that person exists at some time, and at least one other choice in the set where that person never exists. Someone is "non-contingent" relative to a choice set if in every choice in the set, that person (like Eleanor) exists at some time. A "same-people" choice set involves only non-contingent persons. A "contingent" (or in Parfit's terms, "pure") choice set involves only contingent persons. A "mixed" choice set involves some contingent and some non-contingent persons. In a "same-number" choice set, the same number of persons exist no matter what choice we make. In a "different-number" choice set, the number varies with the choice we make.

Observe that in different-number choice sets, at least some people who exist in the choice with the higher number must be contingent. So that leaves us with five categories:

	None contingent	All contingent	Some contingent, some not
Same number	Same people	Contingent same number	Mixed same number
Different number	[Not possible]	Contingent different number	Mixed different number

Our strategy for covering all of these categories proceeds in four steps. We begin with "pairs": choice sets with only two options. First, we assume trade-off rules that tell us which option it is wrong to choose in same-people pairs, insofar as we seek to avoid consigning people to worse fates. Second, we argue that since such trade-off rules are not sensitive to whether people are contingent, the rules that tell us what it is wrong to choose in same-people pairs also tell us what is wrong to choose in contingent and mixed same-number pairs, insofar as we seek to avoid consigning people to worse fates. Third, we propose a way to transform different-number pairs into same-number pairs, so that the rules tell us which option it is wrong to choose in different number pairs as well. Finally, we generalize to choice sets with more than two options by saying that it is wrong to choose an option given that choice set if it would be wrong to choose that option given a pair of that option and another option in the set.

I assume throughout that we have full information about what will happen if we make a choice, including what choices will then be available to us. I assume that a choice A that is available to us in the future, if we make some choice B now, is a choice that we can make now: that is, we can make the composite choice of making B now and A later.

3   Same-people choices

Let's start with same-people choices. They are our theoretical base camp, where our moral judgment feels most at home (or at least as at home as anywhere else).

3.1   Trade-off rules

The first matter to consider is trade offs. Saving some people from worse fates may have the opportunity cost of failing to save other people from worse fates. How should we make such trade offs? We make the "prioritarian" assumption that, other things equal, we have more reason to save people from, as it were, worse worse fates: lower lower levels. And we allow some aggregation. Other things equal, we have more reason to save more people rather than fewer people from a worse fate.

Consider the following choice, in which, if we opt for Most Better, one person drops[3] from 101 to 100, but 99 rise from 101 to 200:

---

[3]Here and throughout I ignore the possibility of status-quo effects: for example, that other things equal there is more reason to keep things as they are than to make a change. The examples should be read in such a way that no option represents the status quo, even if, for convenience, I sometimes use expressions that suggest a status quo, such as "drops."

Most Better   All Equal

Reasonable Trade-Off
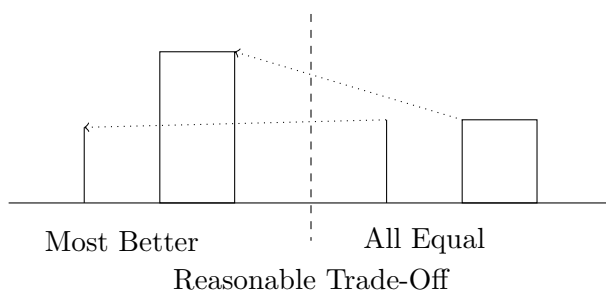
(As usual in such diagrams, the width of the bar represents the size of the population, whereas the height represents how well off the people in that population are. The first "bar" or line in this particular diagram just represents a single person. The dotted arrows indicate the trade offs. If Most Better is chosen, one person is at a lower level, indicated by the downward sloping line, whereas the rest are at a higher level, indicated by the upward sloping line.) Arguably, it isn't wrong to choose Most Better. On the one hand, we give some, but not absolute, priority to reducing (perhaps to zero) the number who occupy lower levels: to saving people from worse worse fates. The fact that All Equal has one fewer at 100 than Most Better has counts more in All Equal's favor than that Most Better has one fewer at 101 than All Equal has counts in Most Better's favor, since 100 is a lower level than 101. On the other hand, we give weight to how many fewer people occupy lower levels: to saving more people rather than fewer from a worse fate. Since Most Better saves so many more from 101 than All Equal saves from 100, this second effect, arguably, swamps the first.

Without presuming to specify what the trade-off rules are, let's assume that there are some trade-off rules such that for any same-people pair $\{A, B\}$, the rules tell us either that A is not wrong or that B is not wrong, given $\{A, B\}$, at least insofar as saving people from worse fates is concerned. This leaves a lot open, of course. But the aim of this paper is not so much to tell us how to treat same-people choices as it is to tell us, once we have decided how to treat same-people choices, to extend that treatment to the other, more exotic choices. So far, then, we have just a kernel of a Worse-Fate principle, which tells us at least when choices are wrong given same-people pairs.

3.2 Zero Embargo

Although, in the examples so far, all of the people in question had lives worth living (represented by positive levels) in every outcome, Worse-Fate and the trade-off rules should also be understood to cover cases where people have lives such that they have reason to prefer never to have been brought into existence (with negative levels): where people have lives, to use a somewhat unfortunate phrase, not for them worth living. For example, it is wrong to choose All Negative, even though the fate for a single person of being at -101 in One Negative is worse than the fate for a single person of being at -100 in All Negative.

6

One Negative   All Negative

Reasonable Trade-Off (Negative)
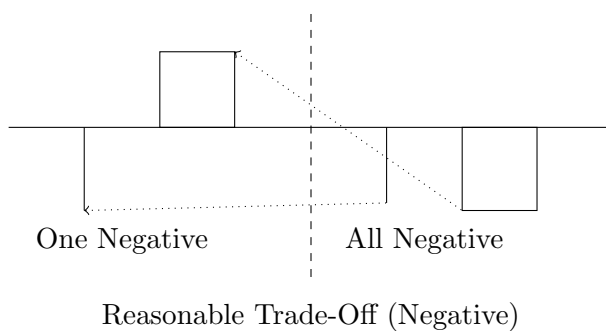
Among the trade-off rules may be Zero Embargo, which gives absolute priority to improving things for people below zero before improving things for people above zero. Zero Embargo can be stated more exactly, to reflect the point that to save from a worse fate is always to reduce the number of people who occupy a lower level, which, in same-people cases, is also to increase the number who occupy a higher level. Zero Embargo gives absolute priority to reducing the numbers at lower levels below zero — to "raising" people below zero closer to zero — over reductions in numbers at lower levels above zero — to "raising" people already at zero or above even higher.

> Zero Embargo: Given the choice of an outcome with exactly n people each at some positive level, it is wrong to choose instead an outcome of $n - m$ people (no matter how many) each at some positive level (no matter how high) and m people (no matter how few) each at some negative level (no matter how close to zero).

Zero Embargo claims that there is a discontinuity at zero, such that gains below zero may not be traded off against gains above zero. There is a ban on exchanges, as it were, across the boundary that divides lives that are not worth living from lives that are. I myself am drawn to Zero Embargo. Why should anyone have to be consigned to a life not worth living just so others, who already have a life worth living, should be even better off? However, I will not assume Zero Embargo in this paper. I will instead indicate where assuming it makes a difference.

3.3 Sufficiency

These trade-off rules may imply a sufficiency threshold or thresholds. That is, the trade-off rules may be such that, first, no choice that leaves everyone above some threshold is wrong. In other words, above that threshold, how we make "trade offs" is not of moral concern. Second, the trade-off rules may be such that any choice that leaves someone below the (same or different) threshold is wrong when the other choice in the pair leaves everyone above that threshold. In effect, Zero Embargo is such a threshold, located at zero.

7

3.4 Equality and "Leveling Down"

The trade-off rules tell us how to triage, when saving some people from worse fates has the opportunity cost of failing to save others from worse fates. These rules do not express a preference for equality as such. Indeed, insofar as the trade-off rules are such rules of triage, they not only do not require, but moreover prohibit, "leveling down." At least, this is so for any case of leveling down such that most will agree (or at least feel defensive about not agreeing) that it should be prohibited, or at least not be required. Given a same-people pair $\{A, B\}$, if B differs from A only in that the number in a higher occupied level H in A is reduced by n and the number at some lower occupied level L in A is increased by n, and all are below the level of sufficiency, then B is wrong. B simply assigns those additional n to the worse fate of L (rather than H), while offering nothing in return as far as saving from worse fates is concerned. This is not meant to rule out the possibility of reasons of a different kind, independent of saving from worse fates, for equalizing as such. Perhaps a more equal state of affairs is, in that respect, impersonally better. I find that hard to credit, though. It seems like a fetish for a certain kind of cosmic pattern. At any rate, it is a different sort of reason.

3.5 Ties

Suppose the trade-off rules leave a tie; more than one outcome is not wrong. However, everyone is below the threshold, if there is one, above which any "trade offs" are not of moral concern. And a lottery is possible. Then we should give each person who might be better off if the tie is resolved in a certain way the highest chance of the tie being resolved in that way, compatible with fairness to others. So if in A, Ada is at 10 and Bill is at 11 and in B, Ada is at 11 and Bill is at 10, then we should decide between A and B by tossing a fair coin. That gives Ada the highest chance of the eleventh unit, namely a 0.5 chance, compatibly with fairness to Bill. To give any higher chance to Ada would mean giving a lower chance to Bill. That would be unfair to Bill, since he is in the same situation. I will view this as not itself part of Worse-Fate, but as an independent Highest Fair Chance Principle which is engaged when Worse-Fate leaves a tie.

3.6 Levels not losses

When evaluating how well one outcome saves people from worse fates relative to the alternatives, I believe, we should consider only the number of people at each level, whoever they may be, at that outcome as compared with the number of people at each level in the alternative outcomes. We should not consider how much better or worse *particular people fare* at that outcome as compared with how *they themselves fare* at some alternative. In a slogan, we should consider "levels not losses." Requiring that our Worse-Fate principle consider levels not losses will have similar consequences to allowing it not to be "narrowly person-affecting," in Parfit's terminology. To be sure, this maneuver—to follow Parfit in

allowing principles not to be narrowly person-affecting—is no great innovation. If there is something new to say here, it is this. While the rejection of a requirement of narrow person-affectingness is usually motivated by exotic choices, in which the choice affects the identities or numbers of people, there is already reason to consider levels not losses even if we consider only garden-variety, same-people choices.

A focus on levels not losses is closely related to what we might call "outcome-anonymity." A principle is outcome anonymous just when once we know, for each outcome, how many people are at which level in that outcome, attaching names—that is, learning who is at which level in which outcome—provides no further information relevant to applying the principle. For example, if in A, Ada is at 10 and Bill is at 11 and in B, Ada is at 11 and Bill is at 10, then either neither outcome should be wrong or both should be, since they differ only in the names attached to the people at the relevant levels. As far as our Worse-Fate principle is concerned, the two outcomes amount to the same thing: one person at 10, one person at 11. Outcome-anonymity blocks sensitivity to losses: to how much better or worse a particular person, such as Ada, is relative to another outcome. For example, outcome-anonymity requires A and B to be equivalent even though A differs from B with respect to, say, how much Ada loses relative to C.

	Ada	Bill
A	10	11
B	11	10
C	13	1

In A she loses 3, in B she loses 2. If outcome-anonymity is independently plausible, then it provides some support for a focus on levels not losses.[4]

A further argument for considering levels not losses is that a principle that considers losses will have difficulty avoiding the following result: that a tie between outcomes that differ only in a permutation of names can be broken by introduction of an outcome that it would be wrong to choose. Suppose that it would be wrong to choose C. Then, one might think, neither A nor B is wrong. Invoking the Highest Fair Chance Principle, we should flip a coin. But now consider a principle that considers losses rather than levels, such as the Minimax Loss Principle, which says that it is wrong to choose an outcome that fails to minimize the maximum difference between someone's level in that outcome and that person's highest level in any outcome (compare Meacham (2012)). According to that principle, the introduction of C, which again is wrong to choose, would make A wrong.

---

[4]That said, the argument for outcome anonymity may seem stronger than it really is if it is mistaken with what we might call "choice anonymity": that changing the names at the top of the columns, reversing Ada's and Bill's levels at every outcome, makes no difference. A principle can be choice anonymous without being outcome anonymous, as is the case with the Minimax Loss Principle presently discussed. So it overstates the case to say that a violation of outcome anonymity is a violation of impartiality or moral equality. For this reason, I worry that Frick (2017, p.356) overstates the case when he labels the axiological analogue of outcome anonymity the "Principle of Impartiality."

This is because in A Ada is 3 below her maximum of 13, whereas in B Bill is only 1 below his maximum of 11.

Why is this result to be avoided? To begin with, there is the intuition that the privileging of B is unfair to Bill. If, when confronted with just A and B, we would think that we should give Bill a 0.5 chance at the eleventh unit, why should the introduction of C lead us to think that giving Bill any chance at the eleventh unit would be wrong? A proponent of the Minimax Loss Principle, presumably, will reply that Ada should be compensated for the additional units she loses by our not choosing C, which would have given her 13. However, this reply violates what we might call the Wrongful Gain Principle: If it would be wrong for someone to have something, then the fact that some choice makes it the case that they go without that thing cannot make that choice wrong. What the master loses in manumitting his slave, what the thief forgoes in not stealing a purse, what the candidate fails to attain in not being the beneficiary of a third party's rigging the election, and so forth are not things for which they are owed compensation. Since Ada would only have the additional units under C if we chose wrongly, the fact that she must be forgo them is likewise not something for which she is owed compensation.

4   Contingent same-number cases

Insofar as Worse-Fate considers levels not losses, it treats contingent same-number cases in precisely the same way as same-people cases. So Worse-Fate implies, intuitively, the deontic equivalent of Parfit (1984)'s "Same Number Quality Claim or Q" at least as restricted to contingent cases: that it is wrong to cause a group of contingent people, the Lows, to live lives at a lower level, still worth living, when we could cause instead the same number of entirely distinct contingent people, the Highs, to live lives at a higher level. (Parfit's "Depletion" and "The 14-Year-Old Girl" are the canonical cases.)

True, if we choose the Lows, we do not make those particular people worse off than they would be if we chose the Highs. But again Worse-Fate considers levels not losses; it is insensitive to how particular people fare relative to another outcome. And, as we saw, there is already reason in same-people cases to favor a principle that considers levels not losses, even before considering exotic cases, in which identities change.

5   Mixed cases

Mixed cases involve both contingent people, whose existence depends on our choice, and non-contingent people, who will exist no matter what we choose. Since Worse-Fate considers levels not losses, it again treats contingent people just like non-contingent people. So it implies that an option in a mixed same-number case is wrong just when the corresponding

option in a corresponding same-people (or contingent same-number) case would be.[5]

This gives us intuitive results in mixed same-number cases. Consider Parfit (2017)'s "Case Four":

	Tom	Dick	Harry
A	50 years of pain		One day of pain
B		One day of pain	Two days of pain

As Parfit observes, a view that favored non-contingent people, such as Harry, perhaps on the ground that Harry would otherwise have been worse off, might make it wrong, counterintuitively, to choose B.

Worse-Fate, by contrast, because it considers levels not losses, implies that it is wrong to make a choice in a mixed case that would be wrong in the corresponding same-people case. To construct such a case, we simply shift cells to the right to fill in any empty cells, like so:

	Dummy	Harry
A'	50 years of pain	One day of pain
B'	One day of pain	Two days of pain

Since A' is wrong and B' isn't given $\{A', B'\}$, A is wrong and B isn't in Case Four.

Having said this, mixed choices may still differ from same-people or contingent choices, because there may be other differences between contingent and non-contingent people that are morally relevant. One such difference concerns the objections that they have to what we choose. If we choose B, then does not non-contingent Harry have a pro tanto objection, on his own behalf, that he could have been better off had we chosen A? By contrast, if we choose A, then contingent Tom has no objection of this kind: no objection that he could have been better off if we had chosen B. In section 15, we will consider a different objection that Tom might still have. But for now let us assume that Tom has no objection. Suppose that the only thing that can answer a person's objection to a choice, so that they do not have a complaint about it, so that we do not wrong them by making it, is that others would have had objections at least as strong to any alternative. It would then follow that we wrong someone by making a choice to which they have an objection when we could have chosen an alternative to which no one would have had any objection. It would then follow that we wrong Harry by choosing B.

However, I think that we should say instead that, insofar as someone's objection to a choice is that they would have been better off had we made a different choice, the fact, if it is a fact, that any other choice would have left people to a worse fate can answer their objection. This is what we can say to Harry when we choose B. The alternative, A,

---

[5]Contrast Voorhoeve and Fleurbaey (2016) and Otsuka (2018). McMahan (2013) explores the possibility that it may depend on what happens in other outcomes on the grounds that it entitles people to regret what might have been.

would have left someone to a worse fate; there would have been someone with 50 years of pain, whereas at B there is only someone with two days of pain. This, it seems to me, reasonably expresses our intuitive disquiet about choosing A, why we feel that Harry's objection has little force in this context This does not deny that there is a morally relevant asymmetry between contingent Tom and non-contingent Harry. Harry has an objection at one outcome, whereas Tom has no objection at any outcome. But this asymmetry makes no difference to the verdict in this case, since Harry's objection can be answered, by citing that the alternative would have meant a worse fate for someone.

It is compatible with this, however, that Harry's non-contingency can be a tie-breaker. Change the case so that Tom in A^*, like Harry in B, will suffer two days of pain. Then choosing B does not save anyone from a worse fate. In that case, perhaps Harry does have an unanswered objection to our choosing B, in which case we wrong Harry by choosing B. Note that this is compatible with the Wrongful Benefits Principle. The gain of another day free of pain in A^* is not something that it would be wrong for Harry to have.[6]

Here's another line of thought to a similar conclusion. Another distinguishing feature of a contingent person, apart from the fact our choosing differently cannot make them better off, is that we cannot *give* a contingent person a chance of existence. We can only *make it the case that* there is a chance that they exist. This may mean that the Highest Fair Chance Principle applies to mixed and same-people cases in different ways. Consider again the choice between A^* and B, which apart from issues of contingency is symmetrical. Should we give Harry only a 0.5 chance of A^*? If we can't give Tom any chance, then Harry may truly say that we can give Harry a better than 0.5 chance of A^*, without thereby reducing the chance we give Tom of B. So perhaps 1 chance of A^* is the highest chance we can give Harry without unfairly depriving anyone else, such as Tom, of a better chance. So, by another route, the fact that Harry is non-contingent is a tie-breaking reason to choose A^* over B.

6   Contingent different-number cases below zero

Now let us consider how to extend Worse-Fate to (contingent and mixed) different-number cases. Before getting lost in the details, let us remind ourselves of the basic strategy. First we transform different-number cases into same-number cases. Then we apply to the transformed cases whatever trade-off rules that we have already decided are appropriate for same-number cases. The transformations are not ad hoc manipulations to deliver intuitive verdicts. They have an underlying rationale. This is to save people from a worse fate, in light of certain observations about what counts as a worse fate: a condition such that in it one would have reason to prefer how things are for people in the alternative. The two most

---

[6]If non-contingency is a tie-breaker, then we should choose Pregnancy Testing over Pre-Conception Testing in Parfit (1984, p. 367), and so resist Parfit's "No-Difference View." Thanks to Andrew Williams for pointing this out.

important observations will be, first, that someone's existing with a life worth living does not save from a worse fate where the alternative is non-existence. Second, non-existence does save from a worse fate where the alternative is someone's existing with a life not worth living.

How we transform different-number cases depends on whether we are considering people in "negative territory," with lives below zero, or people in "non-negative territory," with lives at or above zero. Let's begin by considering cases in which anyone who exists will be in negative territory.

Let's also suppose, for the moment, that at most one person exists. Other things equal, we save people from a worse fate by replacing a person at a lower level, say, -100, with someone at a higher level, say, -50. The best that we can do by way of saving people from a worse fate, however, is simply not to bring anyone into existence. So long as someone comes to be, we are assuming, they will be below zero. So they will have reason to prefer how things are for people in an outcome in which no one exists: "how things are for people" being that no people exist. If the alternative is no one's existing, it is a worse fate to exist below zero.

Now let's suppose that more than one person may exist, while continuing to suppose that anyone who exists is below zero. Bringing one fewer person into existence (by supposition at some negative level) prevents one more person from the worse fate of existing at some negative level.

This licenses the following accounting fiction: treat bringing one fewer person into existence (by supposition at some negative level) as equivalent to replacing a person (at that negative level) with someone at zero. Replacement by someone at zero, like not being brought into existence at all, saves one person (who would otherwise be at some negative level) from a worse fate.

By doing this, we can transform a different-number case into a same-number case. Each fewer person brought into existence becomes, as it were, a person brought into existence at zero. Once we have reduced it to a same-number case, then we can apply whatever trade-off rules we would apply in a same-number case. More generally:

> If outcome A in a pair $\{A, B\}$ has fewer people than outcome B (all of whom, recall, are below zero) add people at zero to A until the groups are the same size, making a new outcome A'. Then apply the trade-off rules to the resulting pair, $\{A', B\}$, to determine which, if either, is wrong. A is wrong given $\{A, B\}$ iff A' is wrong given $\{A', B\}$. B is wrong given $\{A, B\}$ iff B is wrong given $\{A', B\}$.

This explains why we should choose Parfit's "Hell One" with a few below zero over "Hell Two" with vastly many only slightly better off (Parfit 1984, p. 406). The comparison becomes one between Hell Two and, as it were, Hell One$'$, where people are added at zero to equalize the numbers. Assuming plausible trade-off rules, given $\{$Hell One$'$, Hell Two$\}$,

Hell Two is wrong and Hell One$'$ isn't. So given {Hell One, Hell Two}, Hell Two is wrong and Hell One isn't.[7]

7   Contingent different-number cases above zero

Now let's consider cases in which anyone who exists will be in non-negative territory. Let's begin, as before, by supposing that at most one person exists.

In such cases, no one's existing would not save anyone from a worse fate. Existing at some level at or above zero is not a worse fate than not existing. It's not a condition such that someone in it would have reason to prefer how things are for people — namely, not existing — at the alternative. Nor does someone's existing save anyone from a worse fate. Not existing is not a worse fate than existing at a level at or above zero. Indeed, it makes little sense to describe nonexistence as a "condition" such that someone "in that condition" would have reason to prefer existing at a level at or above zero.

This is perhaps the key difference between Worse-Fate and Benefit Thinking. While seeking to prevent people from suffering worse fates would not give us any reason to create people with lives worth living — to "prevent people from not existing" — seeking to benefit people would give us reason to create people with lives worth living, at least if we benefit people by creating them with lives worth living.

Now let's suppose that more than one person may exist, while continuing to suppose that anyone who exists is at or above zero. Given that anyone who exists will be at or above zero, we should not, in contrast to the approach of previous subsection where everyone was in non-negative territory, treat one fewer person at a positive level as equivalent to replacing a person at a positive level with a person at zero. Being at zero rather than at some positive level is a worse fate. But simply not existing rather than being at some positive level, which is what we are actually considering, is not a worse fate.

So I suggest an alternative way of reducing such different-number cases to same-number cases. This is simply to ignore the best-off people in the more populous outcome. First I illustrate how this would work. Then I try to justify it.

> If outcome A in a pair {A, B} has more people than outcome B (all of whom, recall, are at or above zero) subtract from A the best-off people from A until the groups are the same size, making a new outcome A'. Then, as before, apply apply the trade-off rules to the resulting pair, {A', B}, to determine which, if either, is wrong. A is wrong given {A, B} iff A' is wrong given {A', B}. B is wrong given {A, B} iff B is wrong given {A', B}.

---

[7]As Alex Voorhoeve pointed out to me, this accounting fiction — adding a person at zero for each fewer person — is more questionable if we hold that equality itself matters. Hell One$'$ has significant inequality, whereas Hell One does not. While some may be troubled by this, I am not. As I noted in section 3.4, I do not think that equality itself matters.

Worse-Fate thus implies that it is wrong to make the repugnant choice: to bring it about that the Many exist with lives barely worth living, rather than that the Few exist with better lives Parfit (1984).
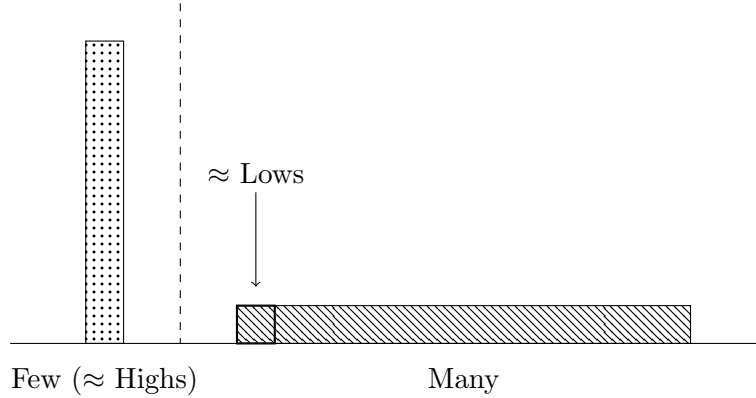
Repugnant Choice

We subtract a sufficient number from the Many (any of whom counts as among the "best off" in Many) so as to equalize numbers with the Few.[8] We are then left with a contingent same-number case like the choice between the Highs and the Lows, where the Few have become the Highs and the Many have become the Lows. We have already agreed that it would then be wrong to choose the Lows over the Highs. So, likewise, it would be wrong to choose the Many over the Few.[9]

Now on to justification. To explain how this approach embodies Worse-Fate Thinking, why it is an appropriate generalization of the single-person case with which we began this subsection, we proceed by cases. As shorthand, say that A "loses" to B iff A is wrong given $\{A, B\}$ and B isn't wrong given $\{A, B\}$.

First, suppose B does not lose, according to the trade-off rules, to A', the reduced outcome with the surplus, best-off population removed. Could B somehow lose to A, the unreduced outcome with the surplus population restored? Or suppose that A' loses to B. Could A somehow not lose to B? (The Repugnant Choice illustrates both possibilities.) In other words, could the unreduced A somehow do better against B than does the reduced A'? It is hard to see how. For A merely adds to A' people at some levels or other. Adding

---

[8]In the diagrams, different fill patterns in the bars indicate different people. So, this diagram represents a contingent different number case, in which none of the people in Many exist in Few.

[9]Another strategy for avoiding the Repugnant Choice is to appeal to the loss of impersonal values (see section 10.2.2) that the Many entail and, perhaps, to deny Negative Transitivity (see section 11.2). Compare Parfit 2016.

people at one level can't reduce the number of people at some other level. And saving people from a worse fate requires reducing the number at some level.

The next sort of case: Suppose that B loses to A', the reduced outcome with the surplus, best-off population removed. And suppose that all of the surplus best off subtracted from A were at or above the highest level in B. Can B somehow not lose to A? Or suppose that A' does not lose to B. Can A somehow lose to B? (Ignoring the Best Off, below, illustrates this second possibility.) In other words, could B somehow do better against the unreduced A than against the reduced A'? Not given our supposition that all of the surplus best off subtracted from A are at or above the highest level in B. B does not save from a worse fate any of the best off "added back" to A, since each person added back is at least as well off as anyone in B.

In the final sort of case, B again loses to A', or A' does not lose to B, but now not all of the best off subtracted from A were at or above the highest level in B. It's less clear what to say about such cases. With an admittedly weaker rationale, and perhaps only as a placeholder, however, I will say that B still can't do better against A. The weaker rationale is that the advocate of A can say, against the advocate of B, "Choose from A any population same in number to B, even the worst off in A, and A does more to save the members of that population from worse fates than does B."

Whether or not we accept this rationale, we can treat this approach as a placeholder for some more nuanced treatment of such cases. None of the cases that we have cause to consider in this paper are of this form. That is, to address classic examples such as the Repugnant Choice, Mere Addition, and so on, we only need to consider cases in which the best off subtracted from A are all better off than the best off in B.

8 Contingent different-number cases above and below zero

To address contingent different-number pairs with people above and below zero, we combine the two approaches. To address choice sets with more than two options, we say that an outcome is wrong given a set if it is wrong when paired with another option in the set. We can now state Worse-Fate, our formalization of Worse-Fate Thinking, in a more general form.

> Worse-Fate: Given a choice set, S, consider each pair of outcomes, $\{A, B\}$ in S. First step: Take whichever outcome, if any, has fewer people below zero. Add dummy zeros to that outcome until the number of people below zero and dummy zeros at that outcome is the same as the number of people below zero in the other outcome. So, suppose that A has no more people below zero than B. Construct A' by adding people at zero to A until the number of people below zero and dummy zeros at A is the same as the number of people below zero in B. Rechristen B as B'. Second step: Having done that, take the resulting outcome, if any, that has more people at or above zero (ignoring dummy zeros).

Subtract the best-off people from that outcome until the number of people at or above zero at the two outcomes is the same. Suppose that B' has no fewer people at or above zero than A'. Construct B'' by subtracting the best-off people in B' until the number of people at or above zero in B'' is the same as the number of people at or above zero in A'. Rechristen A' as A''. Third step: Apply the trade-off rules to the resulting same-number choice set $\{A'', B''\}$ to determine which, if either, is wrong. If and only if A'' loses to B'', then A loses to B (and vice versa). Abstracting from considerations other than how the people in question fare, it is then wrong, given S, to choose an outcome X in S if and only if X loses to some Y in S.

Worse-Fate delivers the right answers in a stylized version of the procreation asymmetry. Given a choice between GOOD, with a child with a life worth living, and NO, with no child, neither choice is wrong. We subtract the child in GOOD so as to make the two populations equal: i.e., equally zero. That leaves us with two equivalent options, neither of which loses to the other. Given a choice between BAD, with a child with a life not worth living, and NO, BAD is wrong. We add a dummy zero to NO so as to make the two populations equal: i.e., equally one. Since it is better to be at zero than below it, BAD loses to NO.

We can put the point another way, to bring out that Worse-Fate doesn't just deliver the right results, but moreover explains the procreation asymmetry. According to Worse-Fate Thinking, it is wrong to leave people to a worse fate, a condition such that they would have reason to prefer how things are for people in some alternative. This is why BAD over NO is wrong. The person who exists in BAD has reason to prefer how things are in NO. By contrast, NO over GOOD is not wrong. No one is left to any fate in NO. There is no person with reason to prefer how things are in GOOD.[10]

The stylized procreation asymmetry assumes a possibly artificially limited choice set, where one cannot instead make, say, a non-contingent child even better off, but not as well off as the child one would have. If one had such a choice, then Worse-Fate implies that, as far as concerns how people fare, one should make it. This does not necessarily mean, however, that it is wrong to have a child when one could instead make a non-contingent child better off. One might have other reasons, such as agent-relative reasons, to have one's own child.

Note that Benefit Thinking, on its own, does not explain why it is wrong to choose to bring someone into existence with a life not worth living when the alternative is not bringing anyone into existence. This is because not bringing someone into existence does not benefit anyone. So Benefit Thinking must be supplemented with a principle that implies that it is wrong to bring someone into existence with a life not worth living when

---

[10]This brings out an important structural similarity to the "Variabilism" of Roberts (2011) and the "SHMV" of Meacham (2012). An important difference, though, is that these alternative approaches consider losses not levels, and so deliver, to my mind, implausible or undermotivated verdicts.

the alternative is not bringing anyone into existence. By contrast, Worse-Fate Thinking gives a unified answer to different number cases below and above zero. In all cases, we seek to avoid consigning people to worse fates.

9 Conflicts with Benefit Thinking

So far, so good for Worse-Fate. But now objections rush in.

9.1 Benefit Requires

Some may object that in the example below, while Worse-Fate permits us to choose the Many over the Few, it does not require us to.
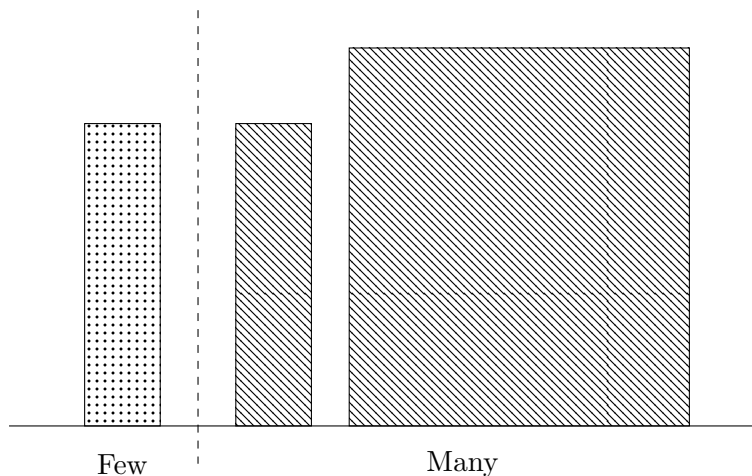
Ignoring the Best Off

Is this acceptable?

The traditional utilitarian answer is no, because is always wrong to fail to bring about a greater sum of happiness. The familiar reply is to deny that we have reason to bring about a greater total quantity of happiness. What matters is making people happy, not making happy people.

A sophisticated proponent of benefit thinking, however, can retrench and say, more plausibly, the following. First, it net benefits a particular person to cause them to be with a life worth living, even if they would not have been worse off otherwise (Parfit 2017). Second, we have reasons, bearing on permissibility, to benefit people, for their sake, not to increase the sum of happiness. In a stronger form, the claim would be:

Benefit Requires: If B net benefits people more than A, then it is wrong to choose A given a choice set that includes A and B.

"Net benefits people more" can be understood in at least two different ways. The first is straightforward:

Sum: B sum net benefits people more than A iff B has a higher sum total of (possibly priority-weighted) net benefits than A.

The second way is something like a generalization of the Pareto criterion familiar from same-people cases: that we can make one person better off without making anyone else worse off. Consider a choice between putting Eleanor in the Good Place or putting Eleanor and Chidi in the Good Place. The latter doesn't benefit Eleanor any less and it additionally benefits Chidi. So it is something like Pareto superior to the former. More generally, we say:

Pareto: B Pareto net benefits people more than A just when we can pair each person in A with exactly one "counterpart" in B, such that every B-counterpart is at least as well off as their A-counterpart, anyone in B without an A-counterpart is above zero, and either there some B-counterpart who is better off than their A counterpart or there is someone in B who has no A counterpart and who is above zero. (More precisely, let $Pop(A)$ be the population at outcome A. B Pareto net benefits people more than A iff there exists an injective "counterpart" function $C : Pop(A) \mapsto Pop(B)$ such that (1) for each $i \in Pop(A), C(i)$ is net benefitted in B at least as much as i in A, (2) for each "counterpartless" $j \in Pop(B)$ (that is, each j such that there is no i such that $j = C(i)$), j is above zero, and (3) either (a) there is some $k \in Pop(A)$ such that $C(k)$ is net benefitted in B more than k in A or (b) there is some counterpartless $j \in Pop(B)$.)

Because a person's counterpart need not be herself, even when she exists in both outcomes, this criterion is in one way weaker than the familiar Pareto criterion. It allows, whereas the familiar criterion does not, that B might Pareto net benefit more than A even when there is someone, Lesser, who is worse off in B than Lesser is in A. This can happen if there is an appropriate counterpart function that assigns Lesser in B a counterpart in A who is not Lesser herself.

There is reason to reject Benefit Requires, on either interpretation. On the sum interpretation, Benefit Requires requires us to make the repugnant choice. After all, the sum of the benefits we give to the Many is greater than the sum of those we give to the Few. And if we accept priority, that reasons to provide increments of net benefit to a person diminish as those increments bring the person to a higher level, then the case only becomes

stronger.[11] On the Pareto interpretation, Benefit Requires requires us to choose GOOD, over NO, in the procreation asymmetry.

9.2 Benefit Permits

There is, however, a weaker version of Benefit Thinking. Consider a choice set containing just Few and Add in the example below (ignoring Equalize for the time being).
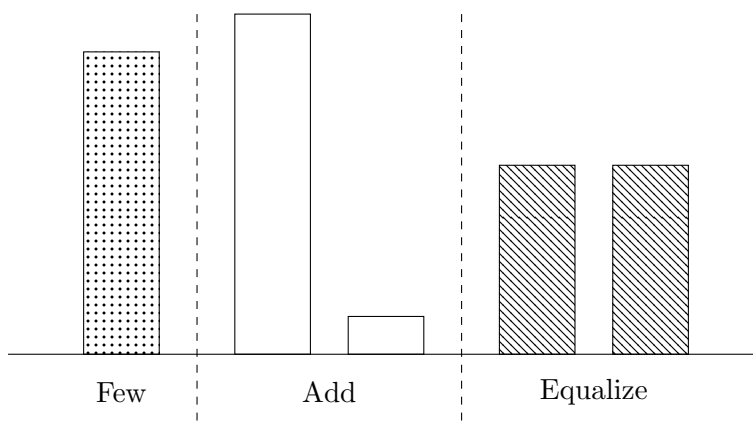
Mere Addition

Once we subtract the better off people in Add, i.e., the first bar, the Few become like the Highs in a contingent same number case, and the worse off people in Add, i.e., the second bar, become like the Lows. Thus, as it is wrong to choose the Lows over the Highs, it is wrong to choose Add over Few.

Is this acceptable? No, one might answer, on the grounds of:

---

[11]Parfit (2017) at footnote 21 seems to suggest that we might avoid the repugnant choice by appealing to a deontic version of Temkin (2012)'s Consolidate Additional Benefits principle: that there is reason to give a few large benefits instead of giving many smaller benefits, even when the sum of benefits to the many is greater and the many are worse off. So there is reason to give large benefits to the Few, instead of smaller benefits to the Many. But, first, it is not clear that CAB applies to the repugnant choice. Temkin (2012) writes that "many find CAB compelling, at least for those cases where if many people have their burdens increased a little this would have relatively little overall impact on their lives, whereas if a few people have their burdens increased substantially this would have a substantial impact on their lives" (pp. 69-70). But the benefit of existence, even a lower level, would seem to have a substantial impact on one's life. Second, CAB may not save us from the repugnant choice. Suppose CAB argues for choosing the Few over the Many in a choice between the two. Nevertheless, CAB should allow us to add n Manys at L while lowering n Few by L (since the same number of benefits, of the same size, go to people with less). So, we ought to take each step down, but once we reach the Many, we ought rather to choose the Few. So, we have cycling. Suppose we then say, as I will say in section 11.4, that, when presented with just the choices in that cycle, any choice is permissible. Then it is permissible to choose the Many from a set that includes the Few.

Benefit Permits: If B net benefits people more than A, then it is not wrong to choose B given $\{A, B\}$.

This is another, weaker expression of Benefit Thinking.

On the sum interpretation, we again have reason to reject Benefit Permits. It permits the repugnant choice.

But what about Pareto Permits: i.e., Benefit Permits interpreted by Pareto? As I argue in section 10, it isn't clear to me that there is a compelling argument, beyond the intuition that Add over Few is not wrong, for any form of Benefit Permits. However, if we like, we can incorporate Pareto Permits into Worse-Fate, by revising the definition of "loses." Recall how Worse-Fate works. It begins by looking at each pair of choices $\{A, B\}$ in the choice set. It transforms each such pair into a same-number choice $\{A'', B''\}$ (by adding zeros to equalize numbers in negative territory and subtracting the best off to equalize numbers in non-negative territory). If and only if A'' is wrong and B'' isn't wrong, we say that A "loses" to B. It is then wrong to choose an X in the set that loses in this sense to some Y in the set. To incorporate Pareto Permits, we change the definition of "loses" to read:

If and only if A'' is wrong and B'' isn't wrong and A does not Pareto net benefit people more than B, A loses to B.

In other words, a choice that Pareto benefits can always, as it were, force a draw. Even if a choice would otherwise lose to another, it doesn't so long as it Pareto benefits. Call the principle that results from Worse Fate by so revising "loses": "Worse-Fate-Pareto."[12]

This might seem ad hoc, but it can be given a rationale. The thought would be that our judgments about ordinary, same-people choices, such as that of Eleanor, are driven simultaneously by two impulses: Worse Fate Thinking and Pareto Benefit Thinking. We are permitted to save from a worse fate, and we are permitted to Pareto benefit. However, we fail to realize that these are different impulses, because in ordinary same-people cases they drive in the same direction. Given any same-number pair, the choice, if there is one, that Pareto net benefits, such as putting Eleanor in the Good Place, will also save people from a worse fate. It's only given exotic, different-number choice sets that a choice that Pareto net benefits may not save people from a worse fate. In such a case, the appropriate extension of our judgments about ordinary, same-people choices is to permit either. That is how we do justice to both impulses: to save from worse fates and to Pareto benefit.

_____

[12]Some might contend that it is wrong to choose Add over Few if a non-contingent person ends up worse off: that is, if there is a non-contingent person in the second bar in Add. To prevent this from counting as a Pareto benefit, we might require the counterpart function to map to oneself when possible: i.e., that $C(i) = i$ where $i \in Pop(B)$. This would be another instance in which non-contingency can break a tie. See the end of section 5.

9.3   The Sadistic Choice

A further worry about Worse-Fate is that it may lead to what Arrhenius (2000) calls a:

> Sadistic Choice: A differs from B only in that A has a group all at or above zero whereas B has a group all below zero, but A loses to B.[13]

The Sadistic Choice is implied by views that hold it against a choice that it adds people above zero but below some positive "critical level" (Blackorby and Donaldson 1984; Blackorby, Bossert, and Donaldson 1995). In the example below, the thick lines represent welfare levels as usual.[14]

---

[13]Arrhenius proves, roughly and translated into deontic terms, that any theory that accepts all four of Negative Transitivity, Dominance, Addition, and Non-Priority Principles cannot avoid permitting the instantiation of at least one of the Repugnant, Anti-Egalitarian, and Sadistic Choices.

- Dominance Principle: If A has the same number as B, and every person in A is better off than any person in B, then A beats B.

- Addition Principle: If when we create B simply by adding to A people all with lower welfare, A beats B, then when we create C simply by adding to A more people with even lower welfare, A beats C.

- Minimal Non-Extreme Priority Principle: There is always some n such that if A has one person below zero and $n-1$ others at very high welfare, whereas B has n people at low but positive welfare, B does not beat A.

- Anti-Egalitarian Choice: A has higher average welfare than B and A is perfectly equal whereas B is unequal, but B beats A.

However, we reject Negative Transitivity and, if we accept Zero Embargo, we reject the Minimal Non-Extreme Priority Principle.

[14]Assume that the people are different, although this isn't indicated by a change of pattern.

Critical Level

B

A
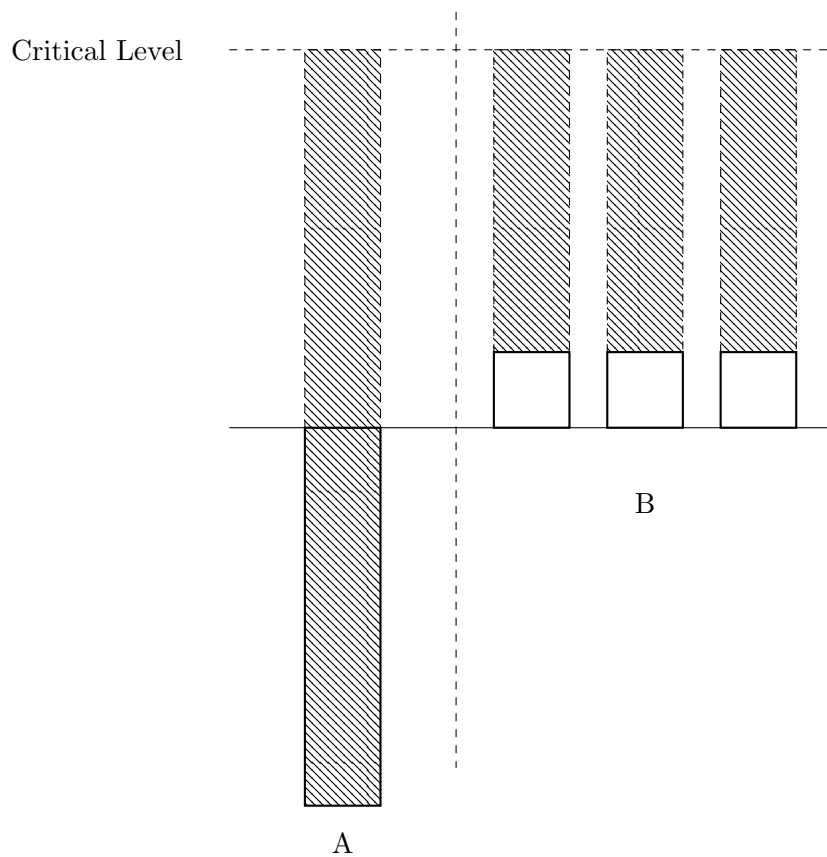
Sadistic Choice: Critical Level

Since the critical level is 50, a rule that tells us to minimize aggregate shortfall would tell us to choose A, which has a total shortfall (shaded) of 100, over B, which has a total shortfall of 120. However, B differs from A only in that A has a group all below zero and B has a group all above zero. (Compare Parfit (1984, p. 415)'s (A).)

Average utilitarianism (and principles, such as Hurka (1983)'s and Ng. (1989)'s, that approximate average utilitarianism at some population levels) also leads to a Sadistic Choice.
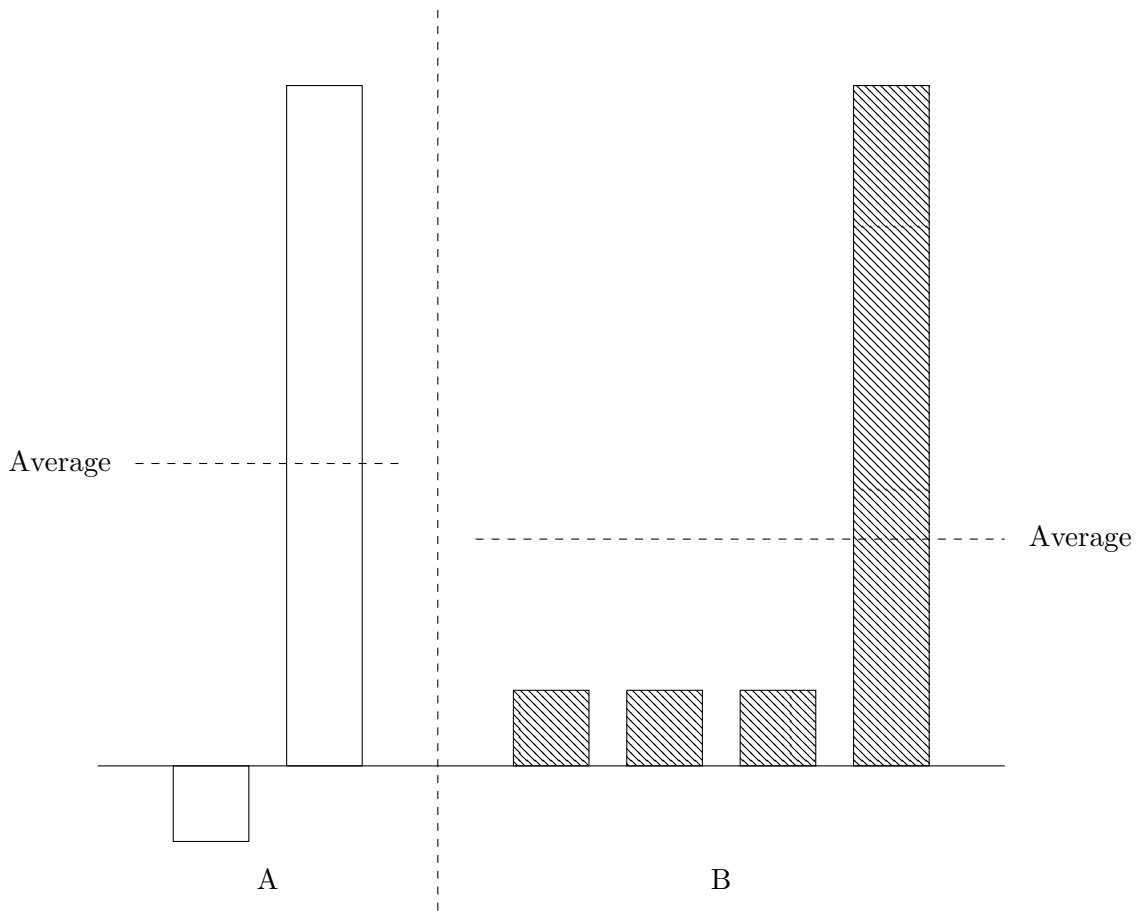
Sadistic Choice: Average Utilitarianism

Since the average in B is lower, average utilitarianism tells us to choose A. However, A differs from B — that is, ignoring the last bar in both — only in that A has a group all below zero and B has a group all above zero.

Can Worse-Fate lead to a sadistic choice? Yes, so long as we reject Zero Embargo.
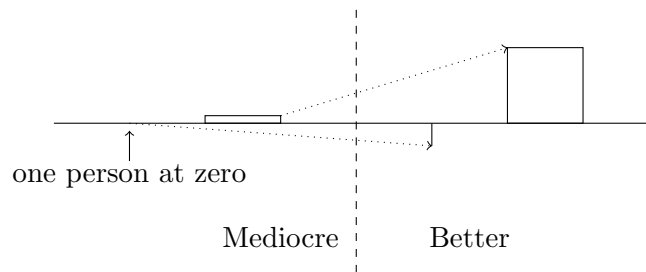
Constructing a Sadistic Choice: Step 1

We start with an outcome, Mediocre, that has just one person at zero and several others all at some equal, positive, Mediocre level. So long as we reject Zero Embargo, we should be able to find an outcome such as Better that beats Mediocre by the trade-off rules. In Better, the person at zero is now slightly negative, and all the rest are at some much higher positive, equal level. By making the second bar of Better sufficiently high, we should eventually be able to make Better beat Mediocre, again so long as we reject Zero Embargo.

Now consider a new choice. Here Mediocre is replaced by Mediocre+ which adds a new population equivalent to the second bar of Better and subtracts the person at zero.
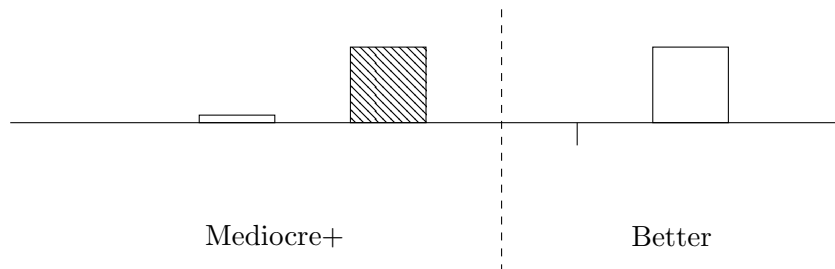
Mediocre+   Better

Constructing a Sadistic Choice: Step 2

If we apply Worse-Fate to this choice, Mediocre+ loses to Better. First, we add a dummy zero to Mediocre+. Second, we subtract the second bar of Mediocre+. In other words, we reduce Mediocre+ to Mediocre. But we have assumed that Mediocre loses to Better. So, according to Worse-Fate, it is wrong to choose Mediocre+ over Better. But the choice of Better over Mediocre+ satisfies the definition of a Sadistic Choice. Mediocre+ differs from Better — that is, ignoring the last bar in both — only in that Better has one person below zero and Mediocre+ has a group all above zero: namely, the first bar.

Should we be troubled by Sadistic Choices of this kind? I don't think so. My sense is that anxiety about Sadistic Choices stems from one of three sources. First, there is the thought that it is clearly not wrong to choose an outcome in which no one is negative over an outcome in which everyone is negative. However, Worse Fate does not imply that it is wrong. In a choice between just the first bar of Mediocre+ and the first bar of Better (that is, the one person below zero), Worse Fate implies that it would be wrong to choose the first bar of Better. Instead, Worse Fate implies that it can be wrong to choose an outcome that differs from another only in that one adds a population in which all are positive whereas the other adds a population in which all are negative. (The latter claim would imply the former claim if we had reason to accept Separation, discussed in section 11.1. But we have independent reason to reject Separation.)

The second source of anxiety is Benefit Thinking. The choice of Better over Mediocre+ is like the choice of Few over Add, in that it passes up an opportunity to Pareto net benefit. Mediocre+ Pareto net benefits people more than Better, just as Add Pareto net benefits people more than Few. What gives the choice of Better over Mediocre+ a "sadistic"

character that the choice of Few over Add lacks is that, without the benefit, one person is below zero. But that may be a red herring. Perhaps the anxiety is just that Mediocre+ loses to Better even though it Pareto net benefits. If this is what troubles one, then one can simply accept Worse-Fate-Pareto.

"No," someone may reply, "the problem is not that we are passing up an opportunity to benefit people. The problem is that we are leaving someone in negative territory, when we have an alternative that leaves no one in negative territory. That's what makes it sadistic, rather than simply inefficient or wasteful." This is the third source of anxiety about Sadistic Choices. I am sympathetic to it. But the more one stresses that leaving someone in negative territory is objectionable in a way in which merely failing to lift someone still further in positive territory is not, the closer one comes to accepting Zero Embargo: that is, that we should not consign people to lives not worth living merely to benefit people with lives worth living. Yet if we accept Zero Embargo, then we avoid Sadistic Choices in the first place. For if we accept Zero Embargo, then it is not wrong to choose Mediocre over Better to begin with.

10   Arguments for Benefit Thinking

We have discussed reasons to reject Benefit Thinking: namely, that, in certain forms, it requires or permits the repugnant choice, or requires GOOD over NO in the procreation asymmetry. If we wish, by opting for Worse-Fate-Pareto, we can incorporate a certain degree of Benefit Thinking into Worse-Fate, without these consequences. But the question remains whether there are any grounds to accept Benefit Thinking. Two main grounds appear time and again in the literature. But, on reflection, they have surprisingly little force.

The first is that only Benefit Thinking can explain the intuitive verdicts in same-number cases. This is simply false. Worse-Fate explains them just as well.

The second ground is that only Benefit Thinking can explain why we have reason to ensure the survival of humanity. This too is false. There are other reasons, albeit ones mostly independent of Worse-Fate.

10.1   Same-number cases

For some time, I (like others, e.g., Ng. (1989, p. 237)) felt compelled to accept not only Benefit Permits, but also Benefit Requires by the following reasoning. "Clearly, in ordinary, same-people cases it is not only not required, but moreover wrong, say, to put Eleanor in the Medium Place, when we could instead put her in the Good Place! How are we to explain this, if not by appeal to Benefit Thinking? The same goes for contingent same-number cases. Doesn't the correct verdict in the choice between the Highs and the Lows require Benefit Thinking?"

But as an argument for Benefit Thinking over Worse-Fate Thinking, this argument has no force. For Worse-Fate on its own explains the correct verdicts. It is wrong to put Eleanor in the Medium Place, or to create the Lows, not because we fail to bestow more benefits, but instead because we leave the worst off to a worse fate.

10.2   The end of humanity

Another argument for Benefit Thinking, which Frick (2017), in criticism, calls the "argument from additional lives," runs as follows: "Not only is it not wrong to postpone the end of humanity, but moreover it is wrong, in some cases, not to postpone it. What reason is there to postpone the end of humanity but to provide future people with the benefit of being caused to be with lives worth living? But once we accept this, aren't we forced to the Benefit principles? Indeed, aren't we pressed even to Benefit Requires and even to the sum interpretation?" (Bennett 1978; Ng. 1989, p. 237; Temkin 2012, p. 414; McMahan 2013, p. 26; Parfit 2017, p. 119).

The reply is that there are other reasons to postpone the end of humanity, besides the benefits that we would thereby confer on future humans.

10.2.1   Interests of non-contingent people

What are these reasons? First, the fact that non-contingent people (mostly those dead or already alive) would be worse off if humanity were to end earlier may argue for postponing the end of humanity. In following Worse-Fate, we will be, to some extent, responsive to such reasons. Postponing the end of humanity may save non-contingent people from a worse fate.

10.2.2   Impersonal values

Second, impersonal values may argue for postponing the end of humanity. The things of impersonal value in question range from the general to the specific: sentient life, valuing, humanity, art and science, and so on. What matters is that the impersonal values have the following four properties.

(1) The things of impersonal value give rise to reasons to ensure that they continue to be instantiated into the future. Perhaps this is because we have reasons to preserve existing things of value (Frick 2017; Scheffler 2018). Perhaps it is because what is of value is, in part, a certain development, or narrative, or career through time, which the premature end of humanity would cut short (Bennett 1978). Perhaps it is simply because the premature end of humanity would prevent things of impersonal value from being instantiated: certain discoveries never made, certain works of art never given shape. (Reasons of this last kind are not only reasons for us to postpone the end of humanity, but also reasons for a creator to create humanity.)

(2) Postponing the end of humanity plays an important, perhaps essential, role in ensuring that these things of impersonal value continue to be instantiated into the future.

(3) These things of impersonal value give us reasons independently of their effects on how anyone fares. Of course, partaking of these impersonal values is what makes people's lives worth living, and so whether these values are available for people to partake of bears on how people fare, which in turn gives us reasons. But these reasons, derived from how people fare, differ from reasons derived from the impersonal values themselves.[15] Indeed, the interests of people provide reasons that weigh against the reasons derived from impersonal values. It may be wrong to purchase a future for those values at the price of human suffering.

(4) The reasons deriving from impersonal values are more holistic and less additive than reasons deriving from the interests of people.

I find it implausible that the reason to be concerned for how a person fares weakens as additional people are added to an outcome. Why should that reason weaken simply because there are more other people, even people distant in time and space? Accordingly, I find it implausible to think that, if we accept that there are reasons to benefit, we can avoid the repugnant choice by saying that, as more people are added, the reason to benefit each additional person diminishes.[16] By contrast, it is entirely plausible that the contribution

---

[15] Discussions of these issues often blur the distinction between interest-based and impersonal-value-based reasons.

> As long as extinction can be deferred, human life, and posthuman life, can continue indefinitely, with unimaginable numbers of people enjoying the goods of life, which might in time become vastly superior to the goods accessible to human beings thus far, just as the goods accessible to us are vastly superior to those that were accessible to our remote evolutionary ancestors. To most of us, it is appalling to think that instead of this incalculable number of people enjoying these incalculable benefits, there might instead be only the emptiness of a world devoid of consciousness (McMahan 2013, p. 26).

Offhand, the indefinite continuation of "human" and "posthuman life," the deferral of "the emptiness of a world devoid of consciousness," would seem to be values independent of people's levels of well-being.

> We can first distinguish between the quality of people's lives and the quantity of well-being per person. These might diverge. The best things in your life might be of a higher quality than the best things in mine, and your life might go worse than mine only because you would have many fewer of these best things (Parfit 2016, p. 117).

On the one hand, Parfit treats "quality," roughly, as a dimension that, multiplied by "duration," gives the "quantity of well-being." On the other hand, he treats "quality" as something that, while still important qua contribution to well-being, independently weighs against the quantity of well-being. But this seems like odd double counting. It seems that when we view quality as something with independent weight, we are not viewing quality as something that is important qua contribution to well-being, but are instead viewing quality as important qua realization of impersonal values. Indeed, that is how Parfit seems to describe it later in the same paper: "There would be no art, or science, no deep loves or friendships, no other achievements, such as that of bringing up our children well, and no morally good people" (p. 123). That would be a loss in itself, presumably, even if other contributions to well-being somehow more than made up for it.

[16] For such an approach to the repugnant choice, see Hurka (1983) and Temkin (2012) on "capped ide-

28

of, say, each additional verse to the flourishing of letters, or each additional painting to the flourishing of the visual arts, diminishes. This is why we do not face a repugnant choice for impersonal values, where we sacrifice a few Dutch Masters for a multitude of velvet Elvises. And there are other holistic, non-additive effects of contributions to impersonal values (some of which are implicit in the very description of the values). Temporal extent matters in a way that spatial extent does not (Bennett 1978; Lenman 2002). Sequence and narrative structure may matter. Continuity may matter.

10.2.3 Agent-relative reasons

Finally, we, the people making the choice, may have various agent-relative reasons to postpone the end of humanity (Bennett 1978; Lenman 2002). Scheffler (2018), in particular, describes "attachment-based" reasons of "interest," "love," "valuation," and "reciprocity." These agent-relative reasons may be closely related to the non-contingent-interest-based and impersonal-value reasons already mentioned. They may simply be reasons based in the interests of non-contingent people, but magnified since we are the non-contingent people. Or what appear to be agent-relative reasons may simply be the reasons that impersonal values give us, in light of what we, in our moment in time, are in a position to affect.[17]

In choices that affect future generations, we should take account of, on the one hand, these reasons of impersonal value and (insofar as they are distinct) these agent-relative reasons, and, on the other, the reasons to save people from worse fates. One possibility is that we should choose in two steps. First, we rule out choices that, by hastening the end of humanity, would be insufficiently responsive to these reasons of impersonal value and these agent-relative reasons. Second, of the choices left, we rule out those that Worse-Fate (or Worse-Fate-Pareto) judges to be wrong. It would then be wrong, or otherwise insufficiently supported by reason, to make any of the choices thus ruled out.

als." In fact, what seems to drive Hurka to claim that increasing the population matters more as the population approaches zero is not the thought that each person's well-being counts for more as the population approaches zero, but instead that each person's existence contributes more to the impersonal value of humanity's existence or survival as the population approaches zero. If so, then his population principle amalgamates two quite different moral concerns: the well-being of persons and "a special value in the existence of animal species" (p. 496).

A further problem for supposing that there is a limit to what Parfit calls the "positive value of quantity" is that if the limit applies to quantity at a given time, then we court Parfit (1984, p. 410-11)'s Absurd Conclusion.

[17]If so, then a Martian who stumbled into our history, without any of our agent-relative reasons, would nevertheless have impersonal reason to help us keep ourselves around long enough, say, to ratify the Equal Rights Amendment.

11 Incoherence

Another source of resistance to Worse-Fate is that it issues patterns of practical verdicts that are said to be somehow "incoherent." In brief, my reply is that we have no good reason to assume that practical verdicts must be "coherent" in the alleged ways. Granted, it would be a problem if accepting Worse-Fate committed us to a logically inconsistent set of beliefs. (We would then know that there is something in Worse-Fate that we should revise, at least if we could identify what it was.) And, granted, if Worse-Fate implies patterns of, say, intransitive practical verdicts, and if we also assume that practical verdicts are transitive, then we are committed to a logically inconsistent set of beliefs. But the question is why we should assume that practical verdicts are transitive in the first place, especially if the intuitively correct practical verdicts pattern in intransitive ways and if we have a principle that explains why they pattern as they do.
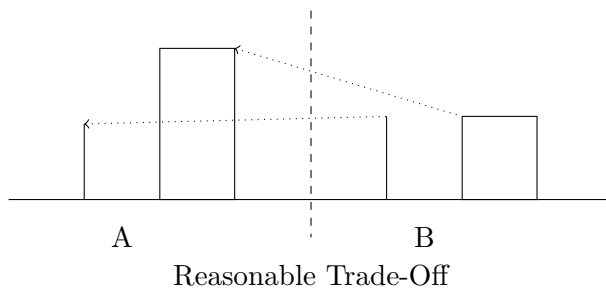
11.1 Separation

Worse-Fate violates:

> Separation: Thinking of outcomes as sets of pairs of persons and levels, where $A' \subseteq A$ and $B' \subseteq B$, if A' is wrong and B' is not given $\{A', B'\}$ and if $B \setminus B'$ is not wrong given $\{A \setminus A', B \setminus B'\}$, then B is not wrong given $\{A, B\}$.

The reasoning behind Separation might be put this way. If B' over A' is a good change, and if $B \setminus B'$ over $A \setminus A'$ is a neutral change, then how can both changes together — B over A — amount to a bad change?[18]

But is Separation compelling? Recall Reasonable Trade-Off.

Reasonable Trade-Off

Let $B' = B$ and $A' =$ the first bar of A. According to Worse-Fate, A' loses to B', $B \setminus B'$ $(= \emptyset)$ does not lose to $A \setminus A'$ $(=$ the second bar of $A)$, but B loses to A. This is a violation of Separation, but anyone who accepts the procreation asymmetry should find this pattern of judgments intuitive.

---

[18] Compare Frick (ms. b)'s argument for why Equalize is not better than Few and Broome (2005)'s puzzlement about the "greediness of neutrality," although the latter concerns, I think, Negative Transitivity, rather than Separation.

11.2 Negative Transitivity

Worse-Fate and Worse-Fate-Pareto also violate:

> Negative Transitivity: If A is not wrong given $\{A, B\}$ and B is not wrong given $\{B, C\}$, then A is not wrong given $\{A, C\}$.

Suppose the outcomes are: NO child, child with a GOOD life, and a child with a GREAT life.[19] According to Worse Fate, GOOD does not lose to NO. It's permissible to have a child with a good life, if the only alternative is not to have a child. That is not for the child to suffer a worse fate. NO does not lose to GREAT. It's permissible not to have a child even if one could have a child with a great life. That is not to suffer any fate at all. But it is not the case that GOOD does not lose to GREAT. It would be wrong to choose to have a child with a good life rather than a great life. In that case, the child does suffer a worse fate.[20]

But why should this be a problem? The intransitive pattern of practical verdicts that Worse-Fate implies is precisely the intuitive pattern! Worse-Fate implies that when both GOOD and GREAT are options, it is wrong to choose GOOD. No other choice from all three, or any two, is wrong.[21]

It is because of the failure of negative transitivity that Worse-Fate-Pareto blocks the "mere addition" argument for the permissibility of the repugnant choice (Parfit 1984; Parfit 2016). Recall the earlier figure:

---

[19]The choice between NO and GOOD is like that facing "Sarah" and the choice between GOOD and GREAT or, perhaps, among NO, GOOD, and GREAT is like that facing "Clare" in Parfit (2017).

[20]Compare (Frick 2020b), which suggests other cases in which, intuitively, Negative Transitivity is violated.

[21]McMahan (2013) appears to use this case to argue for, roughly, Benefit Thinking. If we assume that there is weaker reason for GOOD than for GREAT, and that there is not weaker reason for NO than for GREAT, then it should follow that there is weaker reason for GOOD than for NO. But surely there isn't weaker reason for GOOD than for NO. So, McMahan concludes, there must, in fact, be weaker reason for NO than for GREAT. But I don't see any reason to accept this pattern of judgments. Worse-Fate explains the pattern of judgments that needs to be explained: namely, that any choice is permissible, except GOOD when given the option of GREAT.
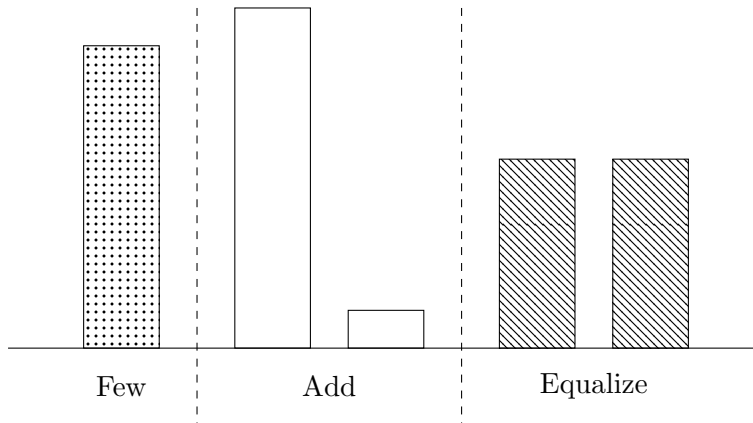
Mere Addition

Suppose that it is not wrong to choose Add over Few. It does not seem wrong to choose Equalize over Add. By repeating the sequence of adding and equalizing, the population increases and the level declines until we reach the Many.

Worse-Fate blocks this argument at the start, by saying that it is wrong to choose Add over Few. Worse-Fate-Pareto blocks it later, by violating Negative Transitivity.

- Given {Equalize, Add}, Equalize is not wrong, since Equalize saves from a worse fate (and Add is wrong because it does not Pareto benefit).

- Given {Add, Few}, Add is not wrong, since Add Pareto benefits.

- Given {Equalize, Few}, Equalize is wrong, since Few saves from a worse fate and Equalize does not Pareto benefit.

This violates Negative Transitivity. Note also that:

- Given {Few, Add, Equalize} only Few is not wrong (since Add loses to Equalize and Equalize loses to Few).

11.3   Choice-Set Independence

Worse-Fate and Worse-Fate-Pareto violate:

Choice-Set Independence: If A is wrong given $\{A, B\}$ then A is wrong given $\{A, B, C\}$.

Begin by noting that there is nothing incoherent, in general, in the idea that the permissibility of a choice can change as the choice set is expanded. For example, consequentialism straightaway violates the converse of Choice-Set Independence, where "not wrong" replaces "wrong." Suppose C brings about a better outcome than A which brings about a better outcome than B. In that case, A is not wrong given $\{A, B\}$ but A is wrong given $\{A, B, C\}$

So why should there be anything troubling about a pattern of verdicts that violates Choice-Set Independence, so long as those verdicts are intuitive? Consider Greaves (ms.)'s illustration of how "necessitarianism"—the view that considers only the well-being of non-contingent people—violates choice-set independence.

	X	Y
A	100	1
B	90	90
C	50	

Given $\{A, B\}$, Y is non-contingent. Since Y's well-being counts along with X's, A is wrong. Given $\{A, B, C\}$, Y is contingent. Since Y's well-being no longer counts, A is not wrong. I agree that this is a compelling counterexample to necessitarianism. But what compels? Opposition to choice-set dependence itself? Or simply recognition that necessitarianism gives the wrong answer—choose A!—given $\{A, B, C\}$? Worse-Fate, it is worth noting, gives the right answer—choose B!—given $\{A, B\}, \{B, C\}$ or $\{A, B, C\}$.

11.4   Acyclicity

Finally, Worse-Fate and Worse-Fate-Pareto violate:

> Acyclicity: If A is not wrong and B is wrong given $\{A, B\}$ and B is not wrong and C is wrong given $\{B, C\}$, then it is not the case that C is not wrong and A is wrong given $\{A, C\}$

It follows that Worse-Fate and Worse-Fate Pareto also violate:

> Positive Transitivity: If A is not wrong and B is wrong given $\{A, B\}$ and B is not wrong and C is wrong given $\{B, C\}$, then A is not wrong and C is wrong given $\{A, C\}$

The example below shows how Worse-Fate violates Acyclicity. A beats B, even though we ignore the first bar of A, and B beats C, since we ignore the first bar of C. But C beats A, since we consider both bars, and the loss in the second bar in C relative to A is, let's suppose, outweighed by the huge gain in the first bar of C relative to A.[22]

---

[22] According to Worse-Fate Pareto, by contrast, B does not beat C, since C Pareto benefits. If we add
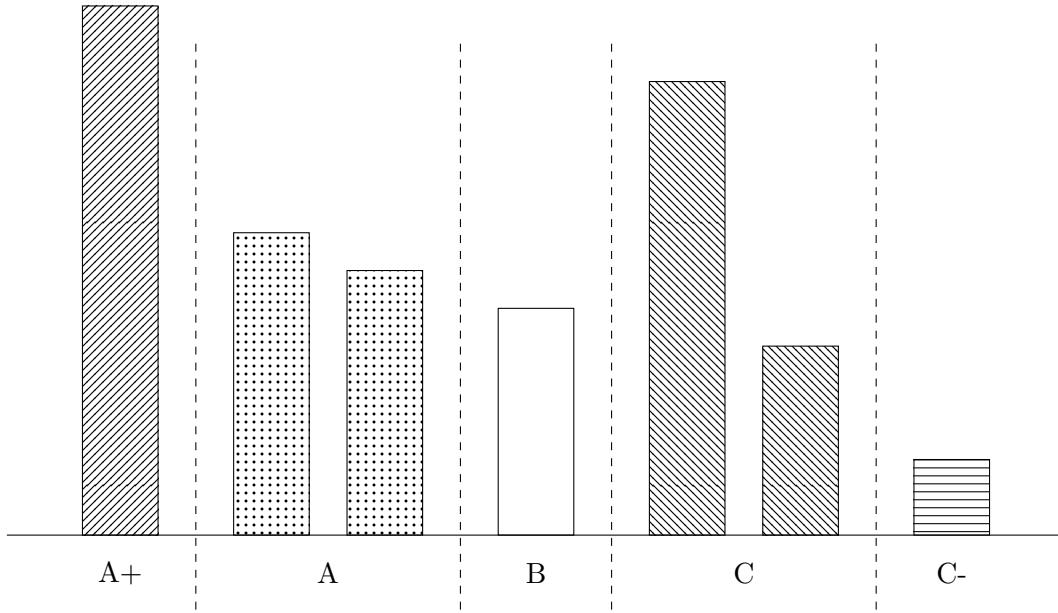
Cycling

How troubling is this? Worse-Fate can still issue verdicts even when such cycles are present, if we add a clause to the final sentence:

> Abstracting from considerations other than how the people in question fare, it is then wrong, given S, to choose an outcome X in S if and only if X loses to some Y in S and it is not the case that Y loses to some Z_1, which loses to some Z_2, which loses to..., which loses to Z_n (all in S), which loses to X.

Given $\{A+, A, B, C, C-\}$, any choice other than $A+$ is wrong. Given $\{A, B, C, C-\}$, only $C-$ is wrong. Given $\{A, B, C\}$, no choice is wrong. Insofar as we are guided by Worse-Fate Thinking, we may well feel that something is wrong with the choice of A, B, or C. Whichever we choose, some alternative would have saved some people from a worse fate. True; it's just that, in light of that fact, the best we can do is just choose one.[23]

_____

the indented clause below, the one that begins "Abstracting from considerations," then C is the only choice given $\{A, B, C\}$ that is not wrong. For a case in which Worse-Fate Pareto violates Acyclicity, consider a case in which the second bar of C is in negative territory but C still beats A by the trade off rules. This is to assume that Zero Embargo does not hold. Since C would "add new people" below zero, C would not Pareto benefit relative to B, and so B would beat C even according to Worse-Fate Pareto. Even if Zero Embargo does hold, Worse-Fate Pareto can violate Positive Transitivity. In Mere Addition, Few beats Equalize, and Equalize beats Add, but Few does not beat Add, since Add Pareto benefits.

[23]As far as I can see, this need not expose us to "money pumps," where we would pay some morally

11.5 The Deontic and the Axiologic

"Incoherence" of the kind that we have been discussing may be problematic for axiological claims, about which outcomes would be better or worse than others. Consider Choice-Set Independence. Whether one outcome is better than another, one might say, should not depend on what other options it is compared with (although Temkin (2012) argues formidably that it does). However, the claims that we are considering are deontic claims, about which outcomes it would be wrong to choose. No one denies that what it is wrong to choose depends on the choices one has.

Of course, if deontic claims about what it would be wrong to choose require an axiological ordering of outcomes as better or worse, then these problems cannot be avoided. And Greaves (ms.) assumes that any moral theory (at least any sane one, that takes the consequences of choices seriously) needs an ordering of outcomes as better or worse. But I don't see why. To be sure, a moral theory needs to say when and why choices can be wrong because of properties of the outcomes that they bring about. But, as Worse-Fate shows, that theory need not say anything about whether one outcome is better or worse than another. (And even if it does, one may suspect that it is simply reporting judgments about which outcomes one ought to choose, given a choice.) Indeed, talk of better and worse outcomes seems at best an unnecessary, and at worst a distorting, intermediate layer. Why say that choosing this outcome is wrong because (i) it is wrong to produce a worse outcome, (ii) this outcome is worse, and (iii) it is worse because of its properties and the properties of alternative outcomes? Why not cut out the middleman and just say that this choice is wrong, because of the properties of the outcome it would produce and of alternative outcomes?

It might be said that we still need an ordering of outcomes to be able to say such things as, "The outcome will be worse if the hurricane makes landfall," where no choice is involved. But I don't see why we need to say such things at all. To be sure, we have reason to hope that the hurricane does not make landfall and to regret it if it does. But an ordering of outcomes is just as superfluous in accounting for reasons for hope and regret as it is in accounting for reasons for action. Why not say that we have reason to hope for this outcome and will have reason to regret the alternative because of the relevant properties of those outcomes: that is, those properties to which we appeal when we, mistakenly in my view, seek to order the outcomes as better or worse?

---

relevant cost x to trade C for B given just those two options, then pay x to trade B for A given just those two options, and then pay x to trade A for C given just those two options. On the assumption of full information, including information about future choices available to us, we can never face such a series of choices. We have all of the choices available to us now: $\{C, B - x, A - 2x, C - 3x\}$. Given that set, every choice but C is wrong.

12   Do we need a principle?

I have proceeded as though we need a principle that tells us how to take account of how future people fare. But it might be said that we need no such principle. Scheffler (2018), for instance, seems to suggest that we can avoid the thicket of population ethics by recognizing simply that the "positive...attachment-independent reasons of beneficence may be...reasons to ensure that the chain of human generations is extended into the indefinite future under conditions conducive to human flourishing."

I'm not sure what he has in mind, but let us suppose that Scheffler is suggesting that so long as a choice leaves everyone above some sufficiency threshold, it is not wrong — at least not wrong because of its effects on how people fare — no matter what the size of the population.

Worse-Fate is certainly compatible with such a view. Indeed, Worse-Fate implies such a view when the trade-off rules reflect a sufficiency threshold.

But worries linger. Scheffler's formula is silent about how we should choose when not everyone can be above the threshold. Such silence is less of a problem, of course, if the threshold is lower and so there is less chance that we will ever be forced to leave people below it. However, if the threshold is higher, there is more chance that we will be forced to leave someone below it and so need guidance that the mere appeal to a threshold of sufficiency does not provide.

If the sufficiency threshold that applies to future people is the same threshold that applies to same-people cases, however, the threshold seems, to judge from intuitions about same-people cases, rather high. What level of healthcare provision for the current population is it, for example, such that trade-offs are no longer a matter of concern?

Now perhaps Scheffler has in mind a sufficiency threshold applying to future people that is different from the sufficiency threshold (if any) applying in same-people cases. But to say that a different threshold applies to future people seems to me hard to defend. Why should the well-being of future people count for less (or more)? People are people.

In any event, the threshold for future people, even if different from the threshold for present people, would also be, intuitively, rather high. Suppose that it costs us nothing at all to excavate ore here rather than there. If we excavate here, as a kind of side-benefit, we prevent some future miasma from being released. If it is released, then people in the future, each of whom enjoys the healthcare, education, etc. of a idealized Scandinavian social democracy, would have to spend a month of afternoons over their lifetime indoors, to avoid the harmless but unbearable unpleasantness until the miasma dissipates. Is it absurd to think that it would be wrong for us, made fully aware of this, to excavate there, rather than here?

13 Ignoring the Unaffected

In a way that we have so far suppressed, the letter of Worse-Fate is at odds with its spirit. The issue is how Worse-Fate counts "unaffected" people. A person is "unaffected" relative to a choice set just when that person is at the same level of well-being at every choice in the set. Those who died long before we make our choice, for example, may be such unaffected people. Of course, any contingent person is affected; their very existence varies with the choice we make. But some non-contingent people may also be affected; although their existence does not vary with our choice, their well-being may. The letter of Worse-Fate makes it sensitive to the unaffected, whereas, according to its spirit, it should be insensitive to the unaffected.

To illustrate what is at stake, consider Parfit 1984, p. 420's objection to average utilitarianism:

> Whether this would be bad, on the Average Principle, depends on facts about all previous lives. If the Ancient Egyptians had a very high quality of life, it is more likely to be bad to have a child now. It is more likely that this child's birth will lower the average quality of the lives that are ever lived. But research in Egyptology cannot be relevant to our decision whether to have children.

Parfit is here assuming a version of average utilitarianism that is sensitive to the unaffected. With respect to the choice whether to have a child now, it takes the average utility not simply of those people who will be affected by that choice, but of all people, including the Ancient Egyptians who are not affected by it. However, a version of average utilitarianism that is sensitive only to the affected would avoid this objection. Since this version of average utilitarianism would not consider the utility of the unaffected Ancient Egyptians in computing the averages, research in Egyptology would have no bearing on our decision whether to have children now.

To tighten this up a bit: For any outcome A in a choice set S, let A^*, the affected counterpart of A, be an outcome that differs from A solely in that every person in A who is unaffected relative to S has been removed. Let S^*, the affected counterpart of S, be the set that replaces each element in S with its affected counterpart. Say that a principle is insensitive to the unaffected just when according to that principle, A is wrong (given S) iff the affected counterpart A^* is wrong (given the affected counterpart S^*).

Read without qualification, many principles are sensitive to the unaffected. First, any principle that gives (sufficient) weight to equality is sensitive to the unaffected. In the following diagram, the unaffected are marked by dashed lines. Whereas A is more egalitarian than B, B^*, with the unaffected removed, is just as egalitarian as A^*.
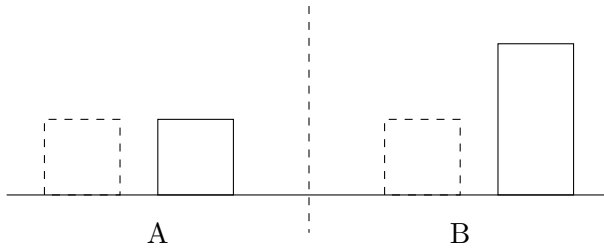
As we have seen with Parfit's Egyptians, average utilitarianism, read without qualification, is also sensitive to the unaffected. Recall the diagram "Sadistic Choice: Average Utilitarianism," in section 9.3. Suppose that the unaffected, marked by dashed lines, are the last bar of each choice. Whereas A has a higher average than B, A^* has a lower average than B^*, where the unaffected are removed.[24]

---
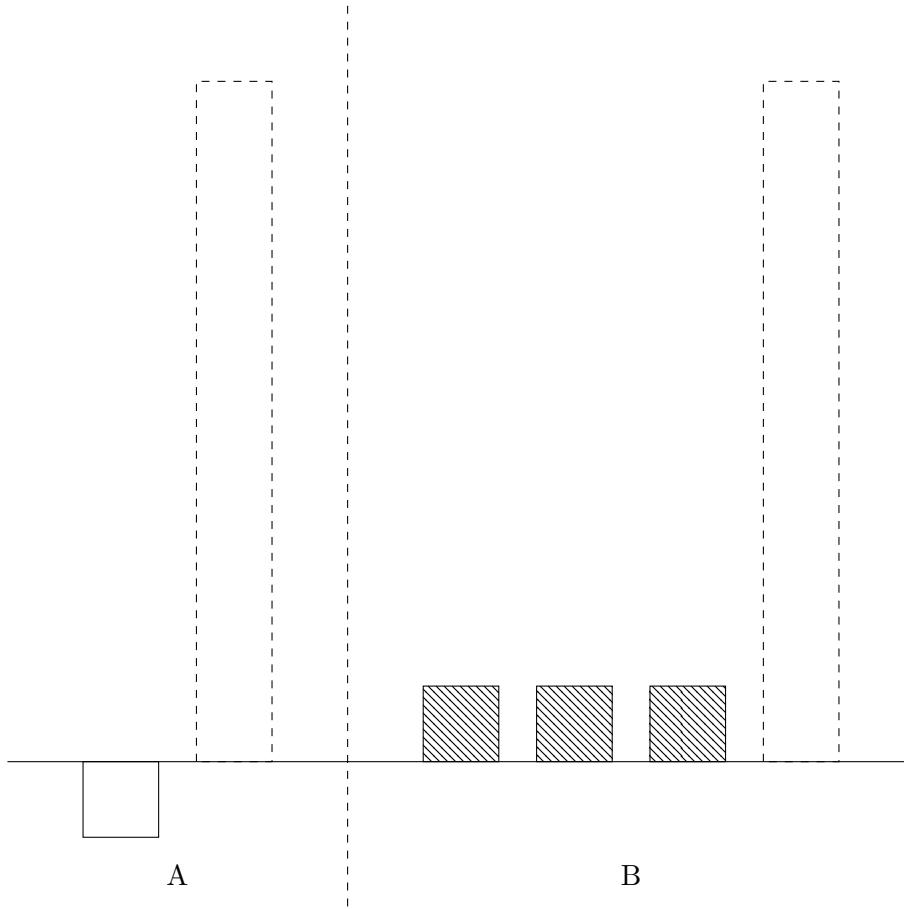
[24]Yet another example of a view that is sensitive to the unaffected is the suggestion in section 10.2.2 that a person's well-being counts for less as the size of the population increases.

As it stands, Worse-Fate is also sensitive to the unaffected. Suppose that the last bar in Mediocre+ and Better, marked by dashed lines, are unaffected. Worse-Fate holds that Mediocre+ is wrong. But Worse-Fate does not hold that the affected counterpart of Mediocre+ — i.e., the last bar removed — is wrong.

Mediocre+   Better
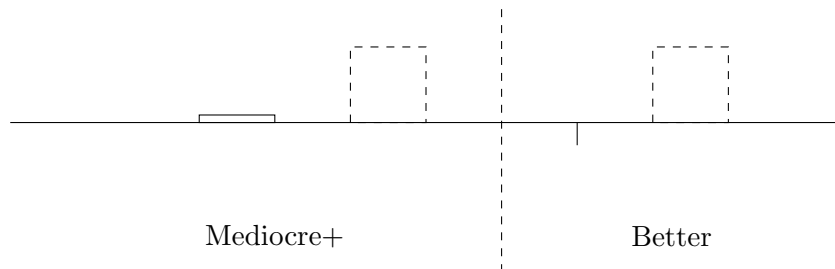
Worse-Fate should not be sensitive to the unaffected. Indeed, Worse-Fate should ignore the unaffected. After all, Worse-Fate cares simply about saving people from worse fates, and no choice can save unaffected people from worse fates. Accordingly, we should modify Worse-Fate so that it judges outcomes only by their affected counterparts. We can now state it in its full generality:

> Worse-Fate[-Pareto] Final Version: Given a choice set, S, consider each pair of outcomes, $\{A, B\}$, in the affected counterpart, S^* of S. Suppose A has no more people below zero than B. Construct A' by adding people at zero to A until the number of people below zero and dummy zeros in A' is the same as the number below zero in B. Rechristen B as B'. Suppose that B has no fewer people at or above zero than A. Construct B'' by subtracting the best-off people in B' until the number of people at or above zero in B'' is the same as the number of people at or above zero in A. Rechristen A' as A''. Then apply the trade-off rules to the resulting same-number choice set $\{A'', B''\}$ to determine which, if either, is wrong. If and only if A'' is wrong and B'' isn't [and A does not Pareto net benefit people more than B] then A loses to B (and vice versa). Abstracting from considerations other than how the people in question fare, it is then wrong, given S, to choose an outcome X in S if and only if X^*, the affected counterpart of X in S^*, loses to some Y^* in S^* and it is not the case that Y^* loses to some Z_1^*, which loses to some Z_2^*, which loses to..., which loses to Z_n^* (all in S^*), which loses to X^*.

Worse-Fate-Pareto differs from Worse-Fate only in including the bracketed clause.

14   Comparison with Frick

It is worth distinguishing this account from that given in recent work by Johann Frick. Frick (2020a) proposes to explain the positive side of the procreation asymmetry by suggesting

that reasons to benefit someone are conditional on the existence of that person. There is no unconditional reason to cause someone to exist, even if they will live a life worth living. However, if someone will exist, then one has reason, for their sake, to benefit them. In choosing NO over GOOD, there is no child for whose sake one fails to benefit. Why, then, isn't it permissible to choose GOOD when one has the options of GREAT and NO? Frick suggests that one will better satisfy the standard of benefitting the child in GREAT for that child's sake, if one chooses GREAT, than one will satisfy the similar standard of benefitting the child in GOOD for that child's sake, if one chooses GOOD. And Frick suggests that, when other things are equal, one has more reason to choose what will better satisfy a relevant standard.

One worry about Frick's account is that it is incomplete. Frick offers no guidance on how it is to generalize. First, it is unclear how it generalizes to contingent same-number pairs where neither choice is (in our sense) Pareto superior to the other (such as a contingent version of Reasonable Trade Off). Second, it is unclear how it generalizes to mixed same-number pairs where, considering just the contingent people, neither choice is Pareto superior to the other. Finally, it is unclear how it generalizes to different number pairs where some in the larger population are worse off than some in the smaller. Consider {Few, Add}. On the one hand, the standard introduced by creating the people in the first bar of Add is better satisfied than the standard introduced by creating the people in the sole bar of Few. On the other hand, the standard introduced by the creation of the people in the second bar of Add is worse satisfied than the standard introduced by the creation of the people in the sole bar of Few. Is Few required, merely permitted, forbidden? For each answer, one can see a plausible argument that it represents the natural extension of what Frick says about the simpler case of {NO, GOOD, GREAT}. (However, Frick (ms. b) suggests that, intuitively, Add should not be forbidden given {Few, Add}, whereas Add should be forbidden and Equalize permitted given {Few, Add, Equalize}. The first judgment does not seem to me obvious, and the second judgment seems to me mistaken.)

Another worry about Frick's account is that its appeal to conditional reasons is unnecessary. We can explain the positive side of the procreation asymmetry by observing, first, that our judgments in ordinary, same-people cases at least as plausibly reflect Worse-Fate as Benefit Thinking, and, second, that given {NO, GOOD}, Worse-Fate Thinking permits NO and permits GOOD, since neither saves anyone from a worse fate. The person who exists in GOOD does not have reason to prefer how things are in NO, and there is no one in NO who has reason to prefer how things are in GOOD.

A further, related worry about Frick's account is that it implies that our reasons with respect to contingent people are different from our reasons with respect to non-contingent people. When people are non-contingent, we have reason to choose for their sake. When people are contingent, we have reason to choose so as to ensure that we will best satisfy whatever standards that will apply to us. If these reasons are so different, then why suppose that we will get the intuitively correct answers in mixed cases: namely, the choices that would be wrong are just those that would be wrong in the corresponding same-people case?

In Case Four, why not suppose that the non-contingent person, Harry, should be favored, so that Tom is consigned to 50 years of pain to save Harry from one day?

15 Who can complain?

Recall the choice whether to create the contingent Lows or the contingent Highs. Each of the Lows has a life worth living. And each would not have existed otherwise. Suppose we accept both:

> No Wrong without Complaint: Abstracting from considerations other than how the people in question fare, it is wrong to choose an outcome only if someone who exists in that outcome can complain about the choice.

> No Complaint when Benefitted: Someone can complain about a choice only if they are not net benefitted by the choice.

Then it is not wrong to choose the Lows over the Highs. And yet it seems that it is wrong, as Worse-Fate implies. So something has to give. This is the "non-identity problem."

Moreover, even if one is not ready to accept these premises in their full generality, one may at least find it intuitive that the Lows could complain about our choice. At least my first reaction to these cases was (and the first reaction of many of my students is) that the Lows do have a complaint. That is, my first reaction was to think: "If we continue to degrade the environment, what will we tell our grandchildren?"

The first point to make is that Benefit Thinking is no better placed to solve the non-identity problem. Granted, a proponent of Benefit Thinking can simply reject No Wrong without Complaint. But then so too can a proponent of Worse-Fate Thinking.

The second point is that Worse-Fate can be part of a more complex view, which sustains the idea that the Lows can complain, whereas the Highs cannot, in part by rejecting No Complaint when Benefitted.

15.1 Complaint although Benefitted

On what basis can someone complain about a choice that net benefits them? One might complain that the choice was ill willed, or negligent, or took excessive risks one's behalf. But when one is caused to be with a life worth living, there needn't be ill will, or negligence, or excessive risk-taking.

Nevertheless, there may be other grounds to complain about being caused to be with a life worth living. One ground for complaint is that the choice may expose you to or cause you to suffer some specific harm (Shiffrin 1999; Harman 2004). A specific harm is a deprivation, bad experience, etc. abstracted from its effect on net well-being. For example,

if we view pain as a specific harm, then causing someone to exist with some pain causes them to suffer a specific harm even if they have a life worth living on balance.[25]

It may suffice to answer this complaint that you consented. But when you are caused to be, you did not consent, because you could not consent. Now, even where you did not consent, because you could not consent, it may suffice to answer this complaint that it saves you from a graver specific harm (Shiffrin 1999). But causing you to be does not save you from a graver specific harm.

Another basis for complaint is that the choice intervenes in something that is yours: something that by all accounts you have a special claim to control. It may suffice to answer this complaint that you consented. And, even where you did not consent, because you could not consent, it may suffice to answer this complaint that it saves you from a graver specific harm. For example, suppose, finding you anesthetized, I see my one chance to surgically remove a fatal tumor. If I wait for your consent, it will be too late. This seems permissible, because the surgery saves you from a graver specific harm. Now suppose that the surgery only improves your vision: saves you from having to replace your glasses with a stronger prescription, for example. While the surgery benefits you overall, it does not save you from a graver specific harm. Do you have no grounds for complaint if I take the initiative? Perhaps you do. And if optical surgery counts as an "intervention," perhaps causing you to be is also an intervention—perhaps a far more significant one?

Of course, you may welcome being benefitted overall (Wallace 2013). So, you may feel conflicted about complaining. Indeed, you may have become attached to what was brought about, in which case the conflict may be particularly acute. However, the fact that you feel conflicted does not mean that I can adequately answer your complaint about what I did.

15.2   Worse-Fate-as-Answer

Suppose, then, that in causing even the Highs to be, they have some such basis for complaint. How might we answer it? We can say that had we not caused the Highs to be, we would have caused other people, the Lows, to be, and the Lows would have suffered a worse fate than the Highs. This seems a sufficient answer to the Highs' complaint. "If it wasn't you, it would have been the Lows, and they would have suffered a worse fate. Whatever complaint you have, their complaint would have been yet stronger."[26]

---

[25]Harman (2004) suggests that it may be wrong to choose B because it involves greater aggregate specific harm than A, even if B also involves greater aggregate specific benefit, and so greater net benefit, than A. But I worry that this will imply that B would lose to A even though each person in B has a life worth living, is better off than every person in A, and suffers less specific harm than anyone in A. This is because, if the population in B is larger than in A, then the aggregate specific harm in B, although much smaller for each person in B than for each person in A, will outweigh the aggregate specific harm in A. Compare Parfit (1984, p. 407)'s Ridiculous Conclusion.

[26]We are assuming that the Highs and the Lows are the only options. There is no option, or no permissible option, of not causing anyone to exist. For reasons why it might be impermissible to cause no one to exist,

After all, this sort of justification seems to answer complaints about causing someone to be with a life not worth living. Suppose we have to choose between causing Bad to be, with a life so bad as not to be worth living, and someone, Worse, with a life still worse. Were Bad to complain, we could tell Bad: "But the alternative would have been even worse for Worse. Whatever complaint you have, Worse would have had an even stronger complaint. What would we have told him?"

If we can answer the Highs' complaint in this way, then this suggests a different explanation of why it is wrong to choose the Lows. It is that we can tell the Highs, but not the Lows: "If we had chosen the alternative, there would have been people, instead of you all, with stronger complaints than you all have. What would we have told them?" This vindicates the thought that the Lows do, as it first seems, have an unanswered complaint about our choice, even though they would not have existed otherwise. "What will we tell our grandchildren?" isn't confused, on this view. It's the right (albeit rhetorical) question to ask.

Consider the following principle:

> Answer: Even if a person has an objection on their own behalf to a choice, they do not have a valid complaint against it — that is, they are not wronged by it — if (i) that choice saves from a worse fate or (ii) avoids objections that are at least as strong. One or the other fact answers their objection.

Clause (i) suffices to answer the complaints of the Highs. But clause (ii) may bind when more than one choice is not wrong according to Worse-Fate. There are two kinds of complaints that might be avoided. Both kinds can be illustrated by Ignoring the Best Off, where neither the Few nor the Many is wrong according to Worse Fate.

First, if there are contingent people among the Few, then we cause someone to be if we choose the Few. Of course, we also cause someone to be if we choose the Many, since the Many has more people than the Few. So whether we choose the Few or the Many, we give someone a complaint about being caused to exist. This very fact, according to (ii), answers the objection of the Many. We can say, to the Many: "No matter what we chose, someone would have had the sort of objection that you have." This would mean that we do not wrong anyone by choosing the Many.

Second, recall from section 5 that non-contingent people, such as Harry, may have a complaint about choices that leave them worse off, which can be a tie-breaking reason. So, returning to Ignoring the Best Off, suppose that (even if there are no contingent people among the Few) there is a non-contingent person among the second, better-off bar of the Many, who also exists, at a lower level, in the Few. That non-contingent person would have an objection about choosing Few, since Few leaves him worse off. Perhaps, then, his objection can answer the objection of a contingent person whom we cause to be by choosing

---

see section 10.2.2.

the Many, by saying: "Had we not made that choice, some particular person would have been worse off."

This gives us the following:

> Worse-Fate-As-Answer (or, alternatively, Worse-Fate-Pareto-As-Answer): Relative to a choice set, S, it wrongs a contingent person x to choose an A where x exists, unless (abstracting from considerations other than how the people in question fare) A is not wrong (relative to S) according to Worse-Fate (or, alternatively, Worse-Fate-Pareto) and for every B in S such that B is not wrong and x does not exist at B, either (a) some y exists at B who does not exist at A or (b) some y who exists at A is worse off in B.[27]

If there is a contingent person among the Few, then clause (a) would mean that (at least as far as Worse-Fate is concerned) it does not wrong anyone among the Many to choose the Many over the Few in Ignoring the Best Off. Similarly, the Pareto version would mean that it does not wrong anyone in Add to choose Add over Few in Mere Addition, so long as there is some contingent person among the Few. If there is a non-contingent person among the second better-off bar in the Many, then clause (b) would mean that it does not wrong anyone among the Many to choose the Many over the Few in Ignoring the Best Off.

Worse-Fate-As-Answer might provide a deeper explanation of Worse-Fate. The root question is how we can justify our choice to the people whom we create. It is a sufficient justification that, had we not created them, we would have created other people, and they would have had stronger objections to our choice. Population ethics can thus be brought entirely within the framework of a morality rooted in objections that people have on their own behalf, according to which a choice is wrong just when there are stronger objections to it than to some alternative.

Worse-Fate-As-Answer might also provide a deeper explanation of the Pareto addendum. Suppose we choose Add over Few where there is some contingent person among the Few. It may be a sufficient justification to the Adds to say the following. "Had we not created you, we would have created someone with complaints like yours. Moreover, one of you can point to how things are even for best off person among the Few, another of you can point to how things are even for next best person among the Few, and so on, and all of you Adds can jointly say that you prefer how things actually are for you."

Even so expanded, Worse-Fate-As-Answer would imply that, in the stylized procreation asymmetry, we wrong the child created by the choice of GOOD. Although GOOD is not wrong given {GOOD, NO}, neither (a) nor (b) is satisfied. NO does not cause anyone else to exist, and NO does not make any particular, non-contingent person worse off.

However, in less stylized versions of the procreation asymmetry, non-contingent people, such as the parents, are made worse off by the choice of NO. And Worse-Fate-As-Answer

---

[27]It's a further question whether it not only wrongs someone, but also is wrong, to choose to A.

explicitly abstracts from agent-relative reasons or impersonal reasons. Perhaps, once those reasons are considered, they can answer the complaints of the child caused to be by GOOD.

That being said, we can of course stipulate cases in which there are no such reasons. Take a case of Ignoring the Better Off in which all of the Few are non-contingent and all of the second-bar of Many are contingent. In such cases, the verdict of Worse-Fate-As-Answer, that we wrong each person in the second bar of the Many by choosing the Many, might seem to have all the sense of the suicide note of Thomas Hardy's Little Father Time: "Done because we are too menny."

References

Arrhenius, G. (2000). "An Impossibility Theorem for Welfarist Axiologies". In: *Economics and Philosophy*, pp. 247–266.

Bennett, J. (1978). "Maximising Happiness". In: *Obligations to Future Generations*. Ed. by R. I. Sikora. Philadelphia: Temple University Press, pp. 61–73.

Blackorby, C., W. Bossert, and D. Donaldson (1995). "Intertemporal Population Ethics: Critical-Level Utilitarian Principles". In: *Econometrica* 63, pp. 1303–1320.

Blackorby, C. and D. Donaldson (1984). "Social Criteria for Evaluating Population Change". In: *Journal of Public Economics* 25, pp. 13–33.

Broome, J. (2005). "Should We Value Population?" In: *Journal of Political Philosophy*.

Frick, J. (2017). "On the Survival of Humanity". In: *Canadian Journal of Philosophy* 47, pp. 344–367.

— (2020a). "Conditional Reasons and the Procreation Asymmetry". In: *Philosophical Perspectives*.

— (2020b). "Future Persons and Victimless Wrongdoing". In: *Worauf es ankommt: Derek Parfits praktische Philosophie in der Diskussion*. Hamburg: Felix Meiner Verlag.

— (ms. b). "Context-Dependent Betterness and the Mere Addition Paradox". In:

Greaves, H. (ms.). "Population Axiology". In:

Harman, E. (2004). "Can We Harm and Benefit in Creating?" In: *Philosophical Perspectives* 18, pp. 89–113.

Hurka, T. (1983). "Value and Population Size". In: *Ethics* 93, pp. 496–507.

Lenman, J. (2002). "On Becoming Extinct". In: *Pacific Philosophical Quarterly* 83, pp. 253–269.

McMahan, J. (2013). "Causing People to Exist and Saving People's Lives". In: *The Journal of Ethics* 17, pp. 5–35.

Meacham, C. (2012). "Person-Affecting Views and Saturating Counterpart Relations". In: *Philosophical Studies* 158, pp. 257–287.

Ng., Y-K. (1989). "What Should We Do about Future Generations?" In: *Economics and Philosophy* 5, pp. 235–253.

Otsuka, M. (2018). "How It Makes a Moral Difference that One is Worse Off than One Could Have Been". In: *Politics, Philosophy, and Economics* 17, pp. 192–215.

Parfit, D. (1984). *Reasons and Persons.* Oxford University Press.

— (2016). "Can We Avoid the Repugnant Conclusion?" In: *Theoria* 82, pp. 110–127.

— (2017). "Future People, the Non-Identity Problem, and Person-Affecting Principles". In: *Philosophy and Public Affairs* 45, pp. 118–157.

Roberts, M. (2011). "The Asymmetry: A Solution". In: *Theoria* 77, pp. 333–367.

Scheffler, Samuel (2018). *Why Worry About Future Generations?* Oxford University Press.

Shiffrin, S. (1999). "Wrongful Life, Procreative Responsibility, and the Significance of Harm". In: *Legal Theory* 5, pp. 117–148.

Temkin, L. (2012). *Rethinking the Good.* Oxford University Press.

Voorhoeve, A. and M. Fleurbaey (2016). "Priority or Equality for Possible People?" In: *Ethics*, pp. 929–54.

Wallace, R. J. (2013). *The View From Here.* Oxford University Press.

— (2019). *The Moral Nexus.* Princeton.