

Parity and Pareto

Abstract

Pareto principles are at the core of ethics and decision theory. The Strong Pareto principle says that if one thing is better than another for someone and at least as good for everyone else, then the one is overall better than the other. But a host of famous figures express it differently, with ‘not worse’ in place of ‘at least as good.’ In the presence of parity (or incommensurability), this results in a strictly stronger Pareto principle, which I call Super-Strong Pareto. Super-Strong Pareto, however, yields cyclic betterness and is therefore false. I point out a number of influential arguments—concerning population ethics, collective action problems, and decision-making in the face of parity and uncertainty—that crucially rely on Super-Strong Pareto and are therefore unsound. I then turn to the most influential argument against the possibility of parity—Broome’s collapsing argument—and argue that it likewise relies on Super-Strong Pareto reasoning and is therefore question-begging. Finally, I turn to the much-neglected question of how to justify Strong Pareto. The answer I arrive at, which emphasizes tie-breaking, yields a striking insight, namely that Super-Strong Pareto amounts to the denial of insensitivity to mild sweetening. That is what makes it problematic in the presence of parity.

Keywords: parity, incommensurability, Pareto, aggregation, dominance

1 Introduction

Pareto principles are at the core of ethics and decision theory. They connect betterness in various underlying respects to overall betterness. The most familiar Pareto principles involve reference to goodness for individuals (or to individual preferences). Two such Pareto principles require that overall (or all things considered) goodness mirror individual goodness in cases of unanimity:

Individual-Wise Weak Pareto: If A is better than B for each individual, then A is overall better than B .

Individual-Wise Pareto Indifference: If A is equally good as B for each individual, then A is overall equally good as B .

These Pareto principles are almost universally accepted. My focus will be on a stronger but still almost universally accepted principle:

Individual-Wise Strong Pareto: If A is better than B for some individual and at least as good for every other, then A is overall better than B .¹

We can also consider Pareto principles—more commonly called dominance principles—which apply to acts and involve reference to how good their outcomes are in various states of the world, i.e. ways things beyond your control might be. And we can consider Pareto principles which involve reference to how good things are along various underlying dimensions of value. I will just state the ‘Strong’ versions of these principles, though they can also come in ‘Weak’ and ‘Indifference’ forms.

State-Wise Strong Pareto: If A is better than B in some state and at least as good in every other, then A is overall better than B .

Dimension-Wise Strong Pareto: If A is better than B along some dimension and at least as good along every other, then A is overall better than B .

So far, so good. But strikingly, many prominent figures state Strong Pareto differently, with ‘not worse than’ in place of ‘at least as good as.’ The list includes Arrow [1977, 220], Sen [1997, 6-7], Harsanyi [1955, 309], Rawls [1999, 58], Gibbard [1982, 403], Kamm [1993, 82], Broome [1997, 84], Samuelson [1954, 387], Buchanan and Tullock [1965, 87], Griffin [1986, 149], G.A. Cohen [1995, 160], Raz [1991, 87], Narens and Skyrms [2020, 3], Pettit [1986, 363], Tsui and Weymark [1997, 245], Temkin [1993, 249], and wikipedia², among others.³

But the slide from ‘at least as good as’ to ‘not worse than’ is innocuous only if we assume comparability, the claim that for any two value-bearers, one is at least as the other (or, more or less equivalently, that any value-bearer is either better than, worse than, or equally good as any other).

¹Strong Pareto is also often stated so as to imply Pareto Indifference (and, of course, Weak Pareto): i.e. as the claim that if A is at least as good as B for every individual then A is at least as good as B overall, and if, moreover, A is better than B for some individual, then A is better than B overall. For ease of exposition, I focus on the simpler version in the main text.

²https://en.wikipedia.org/wiki/Pareto_efficiency. Note that not all of these authors refer to the principle as ‘Strong Pareto.’ Some are talking about Pareto optimality, some are talking generically about the Pareto principle, and Kamm leaves it unnamed.

³By contrast, in §89 of the 1909 appendix to his *Manual of Political Economy*, Pareto [2014] himself states the principle essentially as I have, writing, ‘Let us consider any arbitrary position, and let us suppose that we deviate very slightly from it. . . If, by doing so, the welfare of every individual in the community is increased, the new position is obviously of greater advantage to each of them; and conversely, it will be of less advantage if the welfare of every individual is decreased. The welfare of some individuals may, moreover, *remain constant* without these conclusions being changed’ (emphasis added).

Comparability is controversial. It is rejected by the likes of Parfit [1984], Griffin [1986], Raz [1986], and Chang [2002b], who think that two things can be (in varying terminology) incommensurable, roughly equal, imprecisely equal, or on a par. I'll use the latter term in what follows. If two things are on a par, then neither is better than the other, but nor are they equally good. Hence neither is worse than the other, but also neither is at least as good as the other. ('At least as good' is standardly defined as 'better or equally good.')

The possibility of parity—of two things being on a par—is motivated by the famous 'small improvement argument.' Consider your wedding album and your prized first edition of Descartes' *Meditations*. The former is better than the latter in some (e.g., sentimental) respects but worse in other (e.g., cultural) respects. If they were both in danger and you could only save one of them, you'd be flummoxed. Arguably neither is overall better than the other. But nor are they equally good, since if we mildly sweetened one of them—imagine you remember that there's a \$5 note with one of them—that wouldn't 'break the tie' and lead you to conclude that that one is better overall. Instead, the 'at least as good as' relation is incomplete.⁴

If parity is possible, then the version of Strong Pareto expressed by the many authors mentioned above is strictly stronger, for 'at least as good' entails 'not worse,' but not *vice versa*. Call this stronger Pareto principle:

Individual-Wise Super-Strong Pareto: If A is better than B for some individual and not worse for any other, then A is overall better than B .

And, of course, we can consider State-Wise and Dimension-Wise Super-Strong Pareto principles as well.

Should proponents of parity endorse the Super-Strong Pareto principles? No. For they yield cyclic betterness.⁵ To see this, let x , y , and $y+$ be 'amounts'

⁴This argument relies on the assumption that the 'at least as good as' relation is transitive. Really, the small improvement argument suggests that we must abandon either transitivity or comparability, with most proponents opting for the latter. Note also that we could defend transitivity and comparability by saying that while it is determinate and knowable that either one is better or they're equally good, it is indeterminate and/or unknowable which. See Broome [1997] and Dorr et al. [2023]. The fact that so many prominent authors, like those cited earlier, implicitly treat 'not worse' as equivalent to 'at least as good' is at least a small piece of evidence in favor of comparability.

⁵This point has been noted for particular Super-Strong Pareto principles in different contexts by Parfit [2011, 224], Temkin [2012, 429], Hare [2013, 177], Nebel [2019a,b], Muñoz [2023], and Hedden and Muñoz [forthcoming]. For dissent from the claim that betterness is acyclic, see Rachels [1998] and Temkin [2012].

of value, with $y+$ greater than y but y and $y+$ both on a par with x (i.e. neither greater, nor lesser, nor equal). A , B , and C are alternatives (worlds or actions, say) and 1, 2, and 3 are our respects—either individuals, or states, or dimensions of value, as the case may be.

	1	2	3
A	$y+$	y	x
B	y	x	$y+$
C	x	$y+$	y

Table 1: The Cycle

Super-Strong Pareto entails that A is better than B , since A is better with respect to individual/state/dimension 1 and not worse with respect to 2 and 3. It entails that B is better than C , since B is better with respect to 3 and not worse with respect to 1 and 2. And it entails that C is better than A , since C is better with respect to 2 and not worse with respect to 1 and 3.⁶

So a huge number of authors express Strong Pareto in an overly strong form which, in the presence of parity, yields cyclic betterness. In fairness, some of these authors are tacitly or explicitly assuming comparability and hence rejecting parity, and their arguments do not rely on Super-Strong Pareto as opposed to mere Strong Pareto. For them, it’s just a bookkeeping matter: they should be more careful to avoid sliding from ‘at least as good as’ to ‘not worse than.’

But that is not the end of the story. Several prominent arguments in ethics and decision theory crucially rely on Super-Strong Pareto, as opposed to mere Strong Pareto, and are consequently unsound. These include (§2) Broome’s (2004) greediness objection to the intuition of neutrality, (§3) Barnett’s (2018) ‘No Free Lunch’ argument that all collective action problems involve thresholds, and (§4) Bales et al.’s (2014) argument from their strong competitiveness

⁶This problem also impacts discussion of Pareto efficiency (or Pareto optimality). As it is often defined, a state is Pareto efficient just in case it is impossible to make one person better off without making someone else worse off. Pareto efficiency is widely taken to be insufficient for justice, since a state of affairs in which I have everything and you have nothing is Pareto efficient but unjust. But it is also widely taken to be necessary for justice. But this means that in The Cycle, none of A , B , and C is just, for none is Pareto efficient. One might have thought that the absence of a Pareto efficient state of affairs could only arise in infinite cases, where for any state there is a better one, indeed even one in which *everyone* is better off. It is not implausible that justice is impossible in such infinite cases. But it is implausible that justice is impossible even in finite cases like The Cycle. Better to define a state of affairs as Pareto efficient just in case it is impossible to make one person better off while leaving everyone else at least as well off.

principle to the falsity of the main two theories of decision-making in the face of parity. These are all arguments where the possibility of parity is directly relevant. (The impatient reader may wish to skip some of these case studies.)

Of course, there are arguments against the possibility of parity too. Perhaps the most prominent is Broome's (1997) collapsing argument, intended to show that parity is incompatible with vagueness and should therefore be rejected. But as I explain in §5, that argument likewise relies on a kind of Super-Strong Pareto principle and hence begs the question.

I close in §6 by considering how to motivate Strong Pareto. Strikingly, there have been almost no defenses of Strong Pareto in the literature, and those few either only support Weak Pareto and Pareto Indifference, or they overgeneralize to support Super-Strong Pareto as well. I argue that Strong Pareto is best motivated by appeal to the weighing model of reasons and the notion of tie-breaking. This yields an important insight: Super-Strong Pareto just is sensitivity to mild sweetening, and that explains why it's problematic in the presence of parity.

2 Greed

Many of us have the intuition that creating additional people is typically ethically neutral, in the sense that doing so leaves the world neither better nor worse, at least holding fixed the welfares of already-existing people. This is so, at least, provided that the new lives aren't too bad—then the world would be worse—and perhaps also provided that the new lives aren't too good—then the world would be better. As Narveson [1973, 80] famously puts it, 'We are in favor of making people happy, but neutral about making happy people.'

Perhaps as a consequence of this neutrality, economists are justified in evaluating certain policies in terms of their effects on already-existing people, ignoring the possibility that they would bring new lives into existence [Broome, 2004, 144]. In evaluating projects to improve road safety, or to reallocate health care resources, they take into account how many existing people's lives would be saved, but they rarely take account of the lives that might be added if these saved people themselves had children and grandchildren.

Broome argues against the intuition of neutrality. He first notes that if we analyze neutrality as equally goodness, then there can be only one neutral welfare level—one level such that adding a person with that welfare and holding

everything else fixed leaves the world equally good. If a population with an extra person at a given welfare level is equally good as the population without the extra person, then a population with the extra person at a higher level would be better than the population without the extra person, since it's better (by Individual-Wise Strong Pareto) than the population with the extra person at the lower welfare level. But then the higher welfare level wouldn't be neutral—adding someone with that welfare would make the world better.

He then considers the possibility of analyzing neutrality as parity, so that adding an extra person at a neutral welfare level leaves the world neither better nor worse nor equally good, but rather on a par. This might be motivated by the idea that non-existence is neither better, nor worse, nor equally good for a person as existence with a welfare in the neutral range. But Broome objects that this would make neutrality 'greedy' in a manner antithetical to those who have the intuition of neutrality. Consider the following three populations, where the i^{th} element of each vector represents the welfare of the i^{th} person in that population, 1 and 4 are in the neutral range, and Ω represents non-existence:

A: (4, 4, ..., 4, 6, Ω)

B: (4, 4, ..., 4, 6, 1)

C: (4, 4, ..., 4, 4, 4)

C is better than B, since the same people exist in each, and there's greater welfare and greater equality in C than in B. And A and B are on a par, by the parity interpretation of neutrality. Therefore, C is not worse than A, since it's better than something (namely, B), which is on a par with A.

But Broome [2004, 170] thinks that C ought to be worse than A:

Moving from A to C involves two things. First, the second-last person's wellbeing is reduced from 6 to 4. This is a bad thing.

Second, an extra person is added at level 4. This is a neutral thing.

The net effect of one bad thing and one neutral thing should be bad.

Interpreting neutrality as parity, then, would yield 'a sort of greedy neutrality, which is capable of swallowing up badness or goodness and neutralizing it.'

But anti-greediness—the claim that a good thing plus a neutral thing is good, and that a bad thing plus a neutral thing is bad—on its own yields cyclic betterness. If it's neutral to add someone with a given welfare, it should also be neutral to remove them. After all, removal just is addition seen from the perspective of the other world. Now consider the following three populations:

D: (2, 1, Ω)

E: (Ω , 2, 1)

F: (1, Ω , 2)

By anti-greediness, E should be better than D, since the move from D to E involves one good thing (the increase in person 2's welfare from 1 to 2) and two neutral things (the removal of person 1 and the addition of person 3). And one good thing plus two neutral things should be good. This is already bad enough, conflicting with Broome's own principle of Impartiality, which says that if two populations are related by a permutation of people's welfares, then they're equally good.⁷ But it gets worse. F should be better than E, since the move from E to F likewise involves one good thing and two neutral things. And D should be better than F, since the move from F to D likewise involves one good thing and two neutral things. So we've derived a betterness cycle from anti-greediness plus the assumption that there are multiple neutral welfare levels.

Now, ultimately Broome concludes that there is only one neutral welfare level (though it's vague what it is). So he might see this as more grist for his mill. But I think that proponents of parity should be independently skeptical of anti-greediness, for it amounts to Individual-Wise Super-Strong Pareto, which likewise says that a good thing (one population being better for someone than another) and a neutral thing (one population being neither better nor worse for someone than another) must add up to a good thing (the population's being better overall). And we have already seen that Individual-Wise Super-Strong Pareto can yield cyclic betterness even in fixed-population cases, provided that two outcomes can be on a par for a given individual.

Perhaps Broome would respond that anti-greediness cannot be jettisoned by proponents of the intuition of neutrality, since without anti-greediness, neutrality could not justify economists' practice of only taking into account impacts on already-existing people when evaluating projects and policies. But this practice is unjustified even by Broome's own lights, since he thinks that if there is a neutral range, it's not infinitely wide. Some lives are so bad that their addition

⁷Frick [2017, 356-7] also makes this observation, though his case is more complex. Consider:

(3, 2, 5, Ω)

(3, 2, 4, 4)

(3, 4, 4, 2)

By anti-greediness, the first is better than the second. By anti-greediness again, the third is better than the first, since for the first three people, it has greater welfare and greater equality. But this conflicts with Impartiality, which says that the second and the third are equally good.

makes the world worse, and perhaps some lives are so good that their addition makes the world better. In light of that, neutrality can't justify economists' practice of ignoring the possible new lives that a project or policy might bring into existence, for some of those new lives might be outside the neutral range!

3 Freedom

In many collective action problems, things would be much better if some or all people were to perform a given action than if few or none were to. Examples include climate change, voting, and ending factory farming.

Do you have reason to 'do your part' by biking to work, or showing up to the polls, or ordering vegetarian? Perhaps not. For perhaps the problems are so big, and any individual action so small, that things wouldn't be worse if you were to fail to do your part.

Consider Parfit's [1984] famous case of the harmless torturers. There is a victim hooked up to a torture machine, along with n other people, including you, each of whom has a switch in front of them. Flipping the switch will very slightly increase the voltage going into the victim. But the increment is so small that the victim cannot tell the difference between the case where your switch is flipped from the case where it isn't, no matter how many of the other switches are flipped. So arguably, flipping your switch won't make things worse than refraining, no matter what the others do. And so arguably, you have no reason not to flip your switch.

This argument crucially relies on what Barnett [2018, 5] calls the 'no small improvement' principle. Adapted to the present case, it says:

No Small Improvement: The addition or subtraction of a single [flipped switch] cannot (on its own) make her suffering better or worse.

Barnett gives an ingenious 'No Free Lunch' argument against No Small Improvement. Assuming the argument generalizes, this would entail that all other (finite⁸) cases involve thresholds, where there is some number k such that things would be worse if k -many people performed the relevant action (e.g., flipping) than if only $k - 1$ did.

⁸As Hedden [2020] observes, there are collective action problems with infinitely many people where it is logically impossible for any individual's action to make a difference.

If all cases involve such thresholds, then perhaps you have good consequentialist reasons to do your part. It may be very unlikely that your action would determine whether a threshold is hit, but in that unlikely event, your failing to do your part would make things much worse. And so perhaps doing your part has higher expected value than failing to do so.

Here is Barnett's argument against No Small Improvement. Let there be 10,000 identical victims, one on each step of a huge staircase. For each $1 \leq n \leq 10,000$, the victim on the n^{th} step will experience the voltage corresponding to n switches being flipped. So the victim on the 1st step experiences almost no voltage and almost no pain, the victim on the 10,000th step experiences huge voltage and huge pain, but none can tell the difference between the pain that they'd get on their step from the pain they'd get on the step above or below.

Barnett argues that there must nevertheless be some victim whose pain is worse (albeit not introspectively so) than that of the victim on the step below. For if not, then we could implausibly generate a free lunch. Just let the victim on the top step go down to the bottom, with everyone else shifting up one step. This would be better for someone (namely, the victim originally at the top step) and worse for no one. And that means it would make things better overall. But that's crazy! We can't make things better just by rearranging the victims.

But this argument crucially relies on (Individual-Wise) Super-Strong Pareto. Here is the argument, formulated as a *reductio*. By No Small Improvement, the proposed rearrangement would make one person better off and no one worse off. By Super-Strong Pareto, this means that the rearrangement would make things better overall. But rearranging people on the staircase cannot make things better overall. So by *reductio*, No Small Improvement is false.

But if parity is possible, then Super-Strong Pareto is false, and so the argument is unsound. And Barnett cannot rely just on the weaker Strong Pareto, since that would only enable him to conclude that there must be some victim who is *not at least as well-off* as the victim on the next step down. Here would be the argument: Suppose that every victim is at least as well off as the victim on the next step down. But then, the proposed rearrangement would be better for one person and at least as good for everyone else. By Strong Pareto, the rearrangement would therefore make things better overall. But rearranging people on the staircase cannot make things better overall. So by *reductio*, some victim is not at least as well-off as the victim on the next step down.

So Strong Pareto alone would enable Barnett to conclude that there must be some victim who is *not at least as well off* as the victim on the next step

down. But it would not suffice to enable him to conclude that there must be some victim who is *worse off* than the victim on the next step down. And *that* is what No Small Improvement denies. To get that stronger conclusion, Barnett would need to appeal to Super-Strong Pareto.

With only Strong Pareto, it is a live possibility that each additional switch flipped leaves the victim’s pain on a par with how it would have been otherwise.⁹ This is not terribly plausible in the original version of the harmless torturers case, where the pain varies along a single dimension [Hedden, 2020, Carlson et al., forthcoming]. But we can modify the case to make it a non-threshold one by introducing parity [Hedden, 2020]. As we move up each step of the staircase, let the corresponding victim’s pain increase very slightly in intensity but alternate between burning pain and throbbing pain. And let burning pain of a given intensity be on a par with throbbing pain of very similar intensities. This gives us a non-threshold case where no victim is worse off than the victim on the next step down. And again, to rule out *this* possibility, Barnett would need to appeal to (Individual-Wise) Super-Strong Pareto, but we have already seen that this principle is false if parity is possible.

Suppose there can be such parity-laden and hence non-threshold collective action problems, where things would be much worse if few or none did their part than if all or most did so, but where no individual action can make things worse, instead leaving them merely on a par. (Realistic examples might include ones where individual actions affect the number and identities of future people.) In parity-laden, non-threshold collective action problems, do you have reason to do your part? That depends on the correct theory of rational decision-making in the face of parity. That is the topic of the next section.

4 Competition

The central puzzle for decision-making in the face of parity and uncertainty involves the following case from Hare [2010]:

Opaque Sweetening: There are two opaque boxes. One contains your wedding album (*W*) while the other contains your first edition

⁹To be fair, Barnett is responding to theorists like Nefsky [2017, 2746], who argue that that one single flipping of a switch ‘is simply not enough to *change* [her] level of suffering’ (emphasis added). This claim is stronger than No Small Improvement, for leaving the victim’s suffering on a par with how it would have been otherwise is still a way of changing it, even though it’s not a way of making it worse. Strong Pareto is sufficient to rule out this stronger claim, and so I take Barnett to have given a decisive argument against Nefsky’s view.

copy of the *Meditations* (M). W and M are on a par. Which item is in which box was determined by the toss of a fair coin: If heads, then W is in the left box and M in the right, while if tails, then M is in the left box and W in the right. Also, there is a \$5 note on the right-hand box. Which box should you take?

Perhaps you ought to take the sweetened, right-hand box. After all, you have a reason for taking the sweetened box that is not also a reason for taking the unsweetened box (namely, that you'll get \$5 if you take it), and you have no reason for taking the unsweetened box that is not also a reason for taking the sweetened box. (For instance, that you might get your wedding album is a reason for taking the unsweetened box, but it's also a reason for taking the sweetened box, and similarly for the *Meditations*.)

But perhaps it's permissible to take either. After all, you know that no matter how the coin landed, the contents of the sweetened box are not better than the contents of the unsweetened box, or *vice versa*, since by hypothesis $W+$ (the wedding album plus the \$5) is not better than M , and $M+$ (the *Meditations* plus the \$5) is not better than W .

Hare develops a decision theory, *Prospectism*, which gives the first verdict, and another decision theory, *Deferentialism*, which gives the second. Here, in brief, is how they work. Given an incomplete betterness ordering (a quasi-ordering, in the jargon), we can consider various *coherent completions* thereof. Each coherent completion respects the original ordering's *better than* and *equally good as* relations but eliminates parity by taking each instance of one thing being on a par with another and replacing it with the one's being either better than, worse than, or equally good, as the other. And each of these coherent completions can then be represented by a value function which assigns one thing at least as great a number as another just in case the one is at least as good—according to that coherent completion—as the other.

Prospectism then says that an action is permissible just in case there is some value function representing a coherent completion of the incomplete ordering, relative to which no alternative action has higher expected value.¹⁰

Deferentialism is a bit more complicated. Start by looking at the set of all value functions representing coherent completions of the betterness ordering. Now consider possible regimentations thereof, i.e. subsets whose members all

¹⁰See Hedden [2020] for an argument that Prospectism recommends that you do your part in certain parity-laden, non-threshold collective action problems like the modified version of the harmless torturers case from §3.

agree on the values assigned to some pair of outcomes. For each regimentation, consider all the ways of associating each state with a (possibly different) value function in that regimentation. Each such way of mixing and matching value functions with states will yield a score for each action, which is the sum of the values assigned to the action’s outcome in each state, weighted by the probability of that state. Deferentialism then says that an action is permissible just in case there is some regimentation and some way of mixing and matching value functions from that regimentation such that no alternative action has a higher resultant score. For our purposes, the key thing is simply that that Deferentialism is strictly more permissive than Prospectism, since one way of mixing and matching involves associating the *same* value function with each state.

Bales et al. [2014] argue that both decision theories are false, since they violate a compelling principle of rationality. Their main target is Prospectism, as they disagree with its verdict in Opaque Sweetening. And they endorse a weak principle which is violated by Prospectism but not by Deferentialism.¹¹ But they also endorse a stronger principle which, they briefly note, is violated by both. If this principle is correct, then it’s back to the drawing board.

Bales et al. [2014, 460] say that an action is ‘competitive’ just in case ‘for every way the world could be, its consequences are no worse than the consequences of all alternative actions.’ They then endorse:

Strong Competitiveness: If one or more actions are competitive, and other actions are not competitive, it is rationally required to perform a competitive action.

To see why Prospectism and Deferentialism both violate Strong Competitiveness, consider the following pairwise choice, where states 1, 2, and 3 are equiprobable and where W and $W+$ are each on a par with both M and $M+$, but where $W+$ is better than W and $M+$ is better than M .

	1	2	3
A	$W+$	$M+$	\$0
B	M	W	\$3

Table 2: Only B is Competitive

Here, B is competitive, but A is not, since A is worse than B in state 3. So

¹¹Their Competitiveness principle says that if an action is competitive (i.e. not worse in any state than its alternatives), then it is permissible.

Strong Competitiveness says that B is rationally required. But we can spell out the case so that both Prospectism and Differentialism say that A is permissible. Just let there be a coherent completion of the betterness ordering on which $W+$ and $M+$ are equally good, as are W and M , with $W+$ and $M+$ each better than W and M by \$1.50. This coherent completion can be represented by the value function V with $V(W+) = V(M+) = 4.5$, $V(W) = V(M) = V(\$3) = 3$, and $V(\$0) = 0$. Relative to V , A and B have the same expected value, namely 3, and so Prospectism says that A and B are both permissible. And because, as we saw above, Differentialism is strictly more permissive than Prospectism, it likewise says that A and B are both permissible. Therefore, Prospectism and Differentialism both violate Strong Competitiveness. If the latter is true, both of the former are false. But is it true?

No. In a pairwise choice, one action is competitive while the other isn't just in case the former is better than the latter in one state of the world and worse in none. Assuming that one action is better than another if it's rationally required in a pairwise choice between the two, Strong Competitiveness entails:

State-Wise Super-Strong Pareto: If A is better than B in some state and not worse in any other, then A is overall better than B .

But that is disastrous. It entails that there can be actions A , B , and C such that A is better than B , which is better than C , which is better than A . Just consider the case of The Cycle (Table 1), with 1, 2, and 3 interpreted as states. Strong Competitiveness is false, given the acyclicity of betterness.¹²

5 Collapse

The previous three arguments all suffer from their reliance on Super-Strong Pareto principles which, as we have seen, are stronger than Strong Pareto, and disastrously so in the presence of parity. But perhaps we should just reject parity. Many theorists do. They think that what looks like parity is really just vagueness. When it seems as though one thing is neither better than, nor worse than, nor equally good as another, what's really going on is that it's determinate that one of these three relations holds between them, but it's vague

¹²It is also worth noting the independent implausibility of Strong Competitiveness. Modify the decision problem in Table 2 by reducing the probability of state 3 from $\frac{1}{3}$ to something arbitrarily small and by reducing the payoff of act B in state 3 from \$3 to something arbitrarily small. Strong Competitiveness *still* says that only B is permissible.

or indeterminate which it is. And the small improvement argument certainly looks a lot like a Sorites. So appeal to vagueness can render appeal to parity superfluous [Elson, 2017, Dorr et al., 2023].

Broome [1997] goes further. According to him, not only can appeal to vagueness capture what’s going on when it looks like there’s parity. Worse, vagueness *crowds out* parity. Since we have good independent reason to think that some multidimensional comparatives like *better than* are vague, this means we have good reason to think they don’t give rise to parity.

Take a multidimensional adjective F and one object—the standard—which is intermediate in F -ness. Now consider a Sorites sequence of alternatives which are all the same as each other—but different from the standard—on all but one underlying dimension of F -ness. But as we move up through the sequence, the alternatives get F -er and F -er along that remaining dimension.

For example, F could be ‘intelligent’ and the standard could be someone with moderate levels of spatial, verbal, and emotional intelligence. The alternatives could all have the same level of spatial intelligence, which is higher than that of the standard, and all have the same level of verbal intelligence, which is lower than that of the standard. But they vary in emotional intelligence, with the alternatives near the bottom of the sequence having emotional intelligence much lower than that of the standard and the ones near the top having emotional intelligence much higher than that of the standard.

Plausibly, alternatives near the top of the sequence (call it the top zone) are overall F -er than the standard, while alternatives near the bottom (call it the bottom zone) are overall less F than the standard. If there is parity, there will be some alternatives in the middle of the sequence (call it the middle zone) which are on a par with the standard with respect to F -ness.

But given the assumption that the multidimensional comparative *F -er than* is vague, there can’t be a sharp boundary between the alternatives that are F -er than the standard and those that aren’t. Hence there can’t be a sharp boundary between the top zone and the middle zone.

Now take an alternative that is somewhere in the fuzzy boundary between the top zone and the middle zone. It’s indeterminate whether it is F -er than the standard or just on a par with the standard. But it’s determinate that it is not less F than the standard. After all, it’s nowhere near the bottom zone.

Broome claims that this setup is impossible, for it violates his famous *Collapsing Principle*. And so parity cannot coexist with vagueness. Using ‘false’ to

mean ‘determinately false,’ the principle says:¹³

Collapsing Principle: If it is not false that A is F -er than B but false that B is F -er than A , then A is in fact F -er than B .

To justify this Collapsing Principle, Broome [1997, 74] writes,

If it is false that B is F -er than A , and not false that A is F -er than B , then A has a clear advantage over B in respect of its F -ness. So it must be F -er than B . It takes only the slightest asymmetry to make it the case that one thing is F -er than another.

The Collapsing Principle has generated extensive debate and a host of purported counterexamples [Carlson, 2004, Nebel, 2015]. But we are in a position to formulate a new objection, one which, in my view, gets to the very heart of the matter: The Collapsing Principle relies on Super-Strong Pareto reasoning which, as we have seen, must be rejected by any proponent of parity.

To see this, let us follow Broome in adopting a supervaluationist approach to vagueness. For him, ‘false’ means ‘determinately false,’ which in turn means ‘false on every admissible precisification.’ And when we translate the Collapsing Principle into the language of admissible precisifications, we can see that it is equivalent to what we might call:

Precisification-Wise Super-Strong Pareto: If A is F -er than B on some admissible precisification and not less F on any other, then A is (determinately) overall F -er than B .

To drive home the parallel, let F be ‘good,’ in which case the principle says that if A is better than B on one precisification and not worse on any other, then A is overall better than B .¹⁴

Note that Broome’s argument would not go through with only Precisification-Wise Strong Pareto, which says that if A is F -er than B on some precisification and at least as F on every other, then A is F -er than B . This more plausible

¹³For consistency, I have changed the variables in this statement of the principle and the following justificatory quote.

¹⁴I don’t claim that Precisification-Wise Super-Strong Pareto (aka the Collapsing Principle) yields cyclic F -erness. This is because supervaluationists like Broome hold that determinate truth is truth on every admissible precisification. This claim combined with Precisification-Wise Super-Strong Pareto rules out the very possibility of a cyclic structure like that in Table 1, with the columns 1, 2, and 3 being precisifications rather than individuals, states, or dimensions. But if we were to drop the supervaluationist claim that determinate truth is truth on every admissible precisification, then a structure like that in Table 1 would be possible, and Precisification-Wise Super-Strong Pareto would indeed yield cycles.

principle would only rule out the possibility that it's determinate that A is either F -er than B or equally F , but indeterminate which it is. But what Broome wants to rule out is the possibility that it's determinate that A is either F -er than B or on a par with B with respect to F -ness, but indeterminate which it is. To rule *that* out, Broome needs Precisification-Wise Super-Strong Pareto.¹⁵

So far, we have seen that Super-Strong Pareto principles are problematic in the presence of parity, and we have seen several arguments that crucially rely on them in contexts where parity is directly relevant, including Broome's argument against the very possibility of parity. But what exactly is wrong with the various Super-Strong Pareto principles? And on what basis can we accept the popular Strong Pareto principles but not the Super-Strong ones? The next section gives the answer, one which will enable us to further charge Broome's collapsing argument with begging the question against the proponent of parity.

6 Sweetening Again

Whether individual-wise, state-wise, or dimension-wise, we want to accept Weak Pareto, Pareto Indifference, and Strong Pareto, but not Super-Strong Pareto.¹⁶ On what basis can we do so?

Surprisingly, while Strong Pareto is widely endorsed—Adler [2011, 54] calls it 'one of the most basic building blocks of welfare economics'—there has been

¹⁵Note also that the well-known counterexample to the Collapsing Principle from Carlson [2004] cuts against not only Precisification-Wise Super-Strong Pareto, but also Precisification-Wise Strong Pareto. Suppose that it's indeterminate whether rhetorical skill is a dimension of philosophical quality. Suppose also that you have greater rhetorical skill than me, but that we're equal on all the other dimensions of philosophical quality. According to Carlson, there is a precisification of philosophical quality on which we're equally good philosophers; this is one on which rhetorical skill isn't a dimension of philosophical quality. (Our equal goodness as philosophers then follows from Dimension-Wise Pareto Indifference.) And there is a precisification on which you're a better philosopher than me; this is one on which rhetorical skill is a dimension of philosophical quality. (Your being a better philosopher than me then follows from Dimension-Wise Strong Pareto.) So, according to Carlson, you are a better philosopher than me on one precisification and you are at least as good a philosopher on every other precisification, and yet you're not determinately a better philosopher than I am. This would be a counterexample to even Precisification-Wise Strong Pareto. I am not committed to the latter, for reasons explained in the next footnote. But it is a virtue of my diagnosis of Broome's collapsing argument that it is compatible with Precisification-Wise Strong Pareto and hence with the rejection of Carlson's particular counterexample.

¹⁶What about Precisification-Wise Strong Pareto? In my view, this condition is far less well-motivated than the other Strong Pareto principles. If it is false, then it is a virtue of my reason-based justification of Individual-Wise, State-Wise, and Dimension-Wise Strong Pareto that it does not seem to extend to justify Precisification-Wise Strong Pareto as well. While individuals, states, and dimensions are *loci* of reasons, precisifications are not.

almost no discussion of how to motivate it.^{17,18} One might try to appeal to respect for unanimity. This is how Pareto principles are often justified; indeed, Pareto principles are often described or referred to as principles of unanimity (see e.g., Arrow [1963, 96], Fishburn [1973, 83], Sen [2017, ch. 2]).¹⁹ If two things compare in the same way for each individual, then that should be mirrored in how they compare overall. This could give us individual-wise Pareto principles. And we could extend this to justify state-wise Pareto principles, perhaps by talking about unanimity among possible individuals, each of whom is certain of a single state. And we could extend this to justify dimension-wise Pareto principles, perhaps by talking about unanimity among possible individuals, each of whom monomaniacally cares only about a single dimension of value.

But respect for unanimity would only justify the Weak and Indifference forms of our Pareto principles.²⁰ If A is better than B for some individual and equally good for all others, then A is not unanimously better than B . A is only unanimously *at least as good as* B . But Strong Pareto says A is better. So appeal to unanimity may motivate Weak Pareto and Pareto Indifference, but neither Strong Pareto nor Super-Strong Pareto.

There is another argument for Strong Pareto in the literature, one advanced by Sen [2017, 67], Murphy and Nagel [2002, 50], and Adler [2019, 96-7]. Sen writes that whenever the antecedent of Strong Pareto holds, ‘it is in no one’s interest to be at B rather than at A , and it is in someone’s interest to be at A rather than B ; hence it seems reasonable to say that the society, as an aggregate of the individuals, does prefer A to B .’ Murphy and Nagel [2002, 50] write that

¹⁷Just after this quote, Adler simply asks rhetorically why we wouldn’t also accept Strong Pareto once we’ve accepted Weak Pareto.

¹⁸Strong Pareto says that levelling-down always makes things overall worse. Levelling-down involves bring the welfares of the better-off down closer to the welfares of the worse-off while leaving the latter fixed, thereby reducing inequality. Even egalitarians are loath to say that levelling-down can make things overall better, or even fail to make them worse. At most, they might wish to say that it can make them better in a respect—or along a dimension—namely that of equality. For discussion, see Temkin [1993] and Parfit [1997]. See also Hedden and Muñoz [forthcoming] for a defence of Dimension-Wise Strong Pareto.

¹⁹For discussion of the relation between unanimity and Pareto principles, see Mongin [2016].

²⁰Buchanan and Tullock [1965, 87] attempt to motivate Strong Pareto (which they misstate as Super-Strong Pareto) by appeal to respect for unanimity: ‘A change must be demonstrated to make at least one person in the group ‘better off’ without making any other person ‘worse off’... this means, of course, that a change can be definitely shown to increase ‘total welfare’ only if all persons agree, that is, only if there is unanimous consent of all members of the group.’ (In context, they mean ‘if’ rather than ‘only if.’) By why think that the individuals for whom the two states are equally good will consent to the change? Sure, it’s rationally permissible for them to consent, but it’s also rationally permissible for them not to. And in any case, the same holds for individuals for whom the two states are on a par. So even if their unanimity-based argument for Strong Pareto were sound, it would overgeneralize and support Super-Strong Pareto as well.

‘if at least one person is better off in A than in B , and no one is worse off’ then ‘no one could object to the move from B to A .’ And Adler [2019, 96-7] writes that ‘if some are better off in A than in B , and none are worse off, there’s no conflict... There’s an ethical consideration in favor of A (that some are better off there) and none the other way.’²¹ And so A is better than B .

But this ‘no objection’ argument for Strong Pareto is unsound, for it equally supports Super-Strong Pareto. After all, whenever the antecedent of Super-Strong Pareto is satisfied, it is also the case that it is in no one’s interest to be at B rather than at A (and so no one could object to the move to A), and it is in someone’s interest to be at B rather than at A . Then, by the reasoning advanced by Sen, Murphy and Nagel, and Adler, A should be better than B .

Hedden and Muñoz [forthcoming, 11] consider this issue and attempt to modify this argument so that it only supports Strong Pareto. They say that when A and B are on a par, that doesn’t generate an objection to choosing A (or to choosing B). But it ‘may still count *against* either option’s being overall better; it may instead count in favor of [parity] overall.’

But this doesn’t solve the problem. For return to Strong Pareto. Why doesn’t A and B being equally good for someone count against either option’s being better overall (even though it doesn’t generate an objection to choosing one or the other)? If A ’s being on a par with B for someone counts against A being overall better than B , presumably by counting in favor of a competing alternative— A ’s being overall on a par with B —why doesn’t A ’s being equally good as B for someone similarly count against A being overall better than B by counting in favor of a competing alternative— A ’s being overall equally good as B ? What distinguishes equal goodness from parity in this regard?

Here is the solution: Strong Pareto, but not Super-Strong Pareto, follows from the weighing model of reasons. In a case where individuals’ welfares are all that matter, we can divide up all the reasons for choosing (or desiring, favoring, etc.) A or for choosing B into reasons having to do with individual 1’s welfare, reasons having to do with individual 2’s welfare, and so on. When A is equally good for individual i as B , the reasons having to do with i ’s welfare for choosing A are exactly equally counterbalanced by the reasons having to do with i ’s welfare for choosing B . So when we collect all the individuals for whom A is equally good as B , the reasons having to do with their welfares for choosing A are exactly equally counterbalanced by the reasons having to do

²¹In these quotes, I have changed some variables for the sake of consistency.

with their welfares for choosing B . Add in someone j for whom A is better than B . The reasons having to do with j 's welfare for choosing A are weightier than the reasons having to do with j 's welfare for choosing B . So when we add in all those reasons, they tip the scales in favor of A . Individual-Wise Strong Pareto, then, has to do with tie-breaking, and it follows from the weighing model of reasons. Similarly for State-Wise and Dimension-Wise Strong Pareto.

But Super-Strong Pareto does not follow from the weighing model of reasons. For when A is on a par with B for i , the reasons having to do with i 's welfare for choosing A are *not* exactly equally counterbalanced by the reasons having to do with i 's welfare for choosing B . And so when we add in the reasons having to do with j 's welfare, where A is better than B for j , those reasons don't necessarily tip the scales. For there was no tie to be broken.²²

This yields the following insight: endorsing Super-Strong Pareto amounts to rejecting insensitivity to mild sweetening—the characteristic feature of parity—and that's why it's problematic in the presence of parity. Return to the case of your wedding album and your first edition *Meditations*, and consider the coarse-grained dimension of sentimental-cum-cultural value.²³ They are on a par with respect to this coarse-grained dimension. And if we stipulate that neither can be sold, they are equally good with respect to the dimension of money. So they are overall on a par. Now add \$5 to the wedding album. It is now better than the *Meditations* with respect to the dimension of money and on a par—and hence not worse—with respect to the coarse-grained dimension of sentimental-cum-cultural value. By Dimension-Wise Super-Strong Pareto, it is now overall better than the *Meditations*. But that verdict *just is* the denial of insensitivity to mild sweetening!

To hammer the point home, compare two populations. You're a philosopher in the first and a journalist in the second, while I'm a banker in both. Suppose that being a philosopher is on a par with being a journalist. Plausibly, these populations are on a par overall: they're on a par for you and equally good for me. Now mildly sweeten the second by adding \$5 to my salary. By Individual-Wise Super-Strong Pareto, the mildly sweetened population is now better than

²²One might worry that the weighing model of reasons is itself incompatible with parity, since it assumes that all weights must be comparable with one another. I am inclined to think that we can simply drop that assumption. But even if I am wrong, my overall point stands: Strong Pareto, but not Super-Strong Pareto, follows from the idea that reasons can break ties.

²³This is a *rather* coarse-grained dimension. But coarse-grained dimensions are needed in order for the possibility of parity with respect to a dimension to arise. Chang [2002a, 6] goes so far as to claim that *all* dimensions are somewhat coarse-grained and can be subdivided into further more fine-grained dimensions. For her, value is gunky rather than atomic.

the first. But that's just the denial of insensitivity to mild sweetening.

Or compare two bets. The first gives you your wedding album if heads and a mug if tails, and the second gives you your first edition *Meditations* if heads and that same mug if tails. Plausibly, these bets are on a par: they're on a par in the heads state and equally good in the tails state. Now mildly sweeten the second by adding \$5 to its mug outcome in the tails state. By State-Wise Super-Strong Pareto, that mildly sweetened bet is now better than the first. But that's just the denial of insensitivity to mild sweetening.

Finally, consider a set of admissible precisifications of *F-er than*, all of which rank *A* and *B* on a par with respect to *F*-ness. Now mildly 'sweeten' *A* by adding in a single precisification on which *A* is slightly *F*-er than *B*. By Broome's Collapsing Principle, aka Precisification-Wise Super-Strong Pareto, *A* is now *F*-er than *B simpliciter*. But that's a kind of denial of insensitivity to mild sweetening, and hence begs the question against the proponent of parity!²⁴

²⁴Recall from §5 the justification that Broome [1997, 74] gives for the Collapsing Principle, namely that if *A* is *F*-er than *B* on some precisification and not less *F* than *B* on any other, then there is an asymmetry in favor of *A* vis-à-vis *B* with respect to *F*-ness, and that 'It takes only the *slightest asymmetry* to make it the case that one thing is *F*-er than another' (emphasis added).

References

- Matthew Adler. *Well-being and fair distribution: beyond cost-benefit analysis*. Oxford University Press, 2011.
- Matthew Adler. *Measuring social welfare: An introduction*. Oxford University Press, USA, 2019.
- Kenneth J Arrow. *Social choice and individual values, 2nd edition*. Yale university press, 1963.
- Kenneth J Arrow. Extended sympathy and the possibility of social choice. *The American Economic Review*, 67(1):219–225, 1977.
- Adam Bales, Daniel Cohen, and Toby Handfield. Decision theory for agents with incomplete preferences. *Australasian Journal of Philosophy*, 92(3):453–70, 2014. doi: 10.1080/00048402.2013.843576.
- Zach Barnett. No free lunch: The significance of tiny contributions. *Analysis*, 78(1):3–13, 2018. doi: 10.1093/analys/anx112.
- John Broome. Is incommensurability vagueness? In Ruth Chang, editor, *Incommensurability, Incomparability, and Practical Reason*. Harvard University Press, 1997.
- John Broome. *Weighing Lives*. New York: Oxford University Press, 2004.
- James Buchanan and Gordon Tullock. *The calculus of consent: Logical foundations of constitutional democracy*. University of Michigan Press, 1965. Reprinted as *Selected Works of Gordon Tullock*, V. 2. Liberty Fund Inc. 2004.
- Erik Carlson. Broome’s argument against value incomparability. *Utilitas*, 16(2):220–224, 2004. doi: 10.1017/s0953820804000548.
- Erik Carlson, Magnus Jedenheim-Edling, and Jens Johansson. The significance of tiny contributions : Barnett and beyond. *Utilitas*, forthcoming. Available: <https://philpapers.org/rec/CARTSO-85>.
- Ruth Chang. *Making comparisons count*. Routledge, 2002a.
- Ruth Chang. The possibility of parity. *Ethics*, 112(4):659–688, 2002b.
- G. A. Cohen. The pareto argument for inequality. *Social Philosophy and Policy*, 12(1):160–185, 1995. doi: 10.1017/s026505250000460x.
- Cian Dorr, Jacob M. Nebel, and Jake Zuehl. The case for comparability. *Noûs*, 57(2):414–453, 2023. doi: 10.1111/nous.12407.
- Luke Elson. Incommensurability as vagueness: A burden-shifting argument. *Theoria*, 83(4):341–363, 2017. doi: 10.1111/theo.12129.
- Peter C Fishburn. *The theory of social choice*. Princeton University Press, 1973.

- Johann Frick. On the survival of humanity. *Canadian Journal of Philosophy*, 47(2-3):344–367, 2017. doi: 10.1080/00455091.2017.1301764.
- Allan Gibbard. The prospective pareto principle and equity of access to health care. *The Milbank Memorial Fund Quarterly. Health and Society*, pages 399–428, 1982.
- James Griffin. *Well-being: Its meaning, measurement and moral importance*. Clarendon press, 1986.
- Caspar Hare. Take the sugar. *Analysis*, 70(2):237–247, 2010. doi: 10.1093/analys/anp174.
- Caspar Hare. *The limits of kindness*. Oxford University Press, 2013.
- John C Harsanyi. Cardinal welfare, individualistic ethics, and interpersonal comparisons of utility. *Journal of political economy*, 63(4):309–321, 1955.
- Brian Hedden. Consequentialism and collective action. *Ethics*, 130(4):530–554, 2020. doi: 10.1086/708535.
- Brian Hedden and Daniel Muñoz. Dimensions of value. *Noûs*, forthcoming. doi: 10.1111/nous.12454.
- Frances Myrna Kamm. *Morality, mortality: Death and whom to save from it*, volume 1. Oxford University Press on Demand, 1993.
- Philippe Mongin. Spurious unanimity and the pareto principle. *Economics & Philosophy*, 32(3):511–532, 2016.
- Daniel Muñoz. Sources of transitivity. *Economics and Philosophy*, 39(2):285–306, 2023. doi: 10.1017/s0266267122000074.
- Liam Murphy and Thomas Nagel. *The myth of ownership: Taxes and justice*. Oxford University Press, 2002.
- Louis Narens and Brian Skyrms. *The Pursuit of Happiness: Philosophical and Psychological Foundations of Utility*. Oxford University Press, 2020.
- Jan Narveson. Moral problems of population. *The Monist*, 57(1):62–86, 1973. doi: 10.5840/monist197357134.
- Jacob M. Nebel. *Incommensurability in Population Ethics*. PhD thesis, University of Oxford, 2015. Available: <https://philpapers.org/archive/NEBIIP.pdf>.
- Jacob M. Nebel. An intrapersonal addition paradox. *Ethics*, 129(2):309–343, 2019a. doi: 10.1086/700085.
- Jacob M. Nebel. Asymmetries in the value of existence. *Philosophical Perspectives*, 33(1):126–145, 2019b. doi: 10.1111/phpe.12128.

- Julia Nefsky. How you can help, without making a difference. *Philosophical Studies*, 174(11):2743–2767, 2017. doi: 10.1007/s11098-016-0808-y.
- Vilfredo Pareto. *Manual of political economy: a critical and variorum edition*. OUP Oxford, 2014.
- Derek Parfit. *Reasons and persons*. OUP Oxford, 1984.
- Derek Parfit. Equality and priority. *Ratio*, 10(3):202–221, 1997. doi: 10.1111/1467-9329.00041.
- Derek Parfit. *On what matters*, volume 2. Oxford University Press, 2011.
- Philip Pettit. Free riding and foul dealing. *Journal of Philosophy*, 83(7):361–379, 1986.
- Stuart Rachels. Counterexamples to the transitivity of better than. *Australasian Journal of Philosophy*, 76(1):71–83, 1998. doi: 10.1080/00048409812348201.
- John Rawls. *A theory of justice: Revised Ed.* Harvard University Press, 1999.
- Joseph Raz. *The morality of freedom*. Clarendon Press, 1986.
- Joseph Raz. Mixing values. *Aristotelian Society Supplementary Volume*, 65(1): 83–118, 1991. doi: 10.1093/aristoteliansupp/65.1.83.
- Paul A Samuelson. The pure theory of public expenditure. *The review of economics and statistics*, 36(4):387–389, 1954.
- Amartya Sen. *On economic inequality*. Oxford university press, 1997.
- Amartya Sen. *Collective Choice and Social Welfare*. Harvard University Press, expanded edition, 2017. ISBN 978-0-674-97160-8.
- Larry S Temkin. *Inequality*. Oxford University Press, 1993.
- Larry S Temkin. *Rethinking the good: Moral ideals and the nature of practical reasoning*. Oxford University Press, 2012.
- Kai-yuen Tsui and John A Weymark. Social welfare orderings for ratio-scale measurable utilities. *Economic Theory*, 10:241–256, 1997.