# PE3101P: Decision and Social Choice

Lecture Notes 8 — Egalitarianism and prioritarianism

Instructor: Tomi Francis

8 October 2026

### 0.1 Gambles and prospects

I have to correct a verbal mistake from last week. I often used the word "prospect" informally to denote an assignment of states of nature to social outcomes. Officially, though, as in the decision theory half, prospects only encode information about probabilities and not about states of nature. I really meant to say *social gambles*, which are assignments from states of nature to social outcomes.

So, the situation is like this:

| | Takes in | Spits out |
|---|---|---|
| Social gamble | a state of nature | an outcome |
| Social prospect | an outcome | a probability (a number between 0 and 1) |

To move from a social gamble to a social prospect: for each outcome, add up the probabilities of all the states of nature resulting in that outcome. This is the probability of the outcome. Figure 1 shows an example.

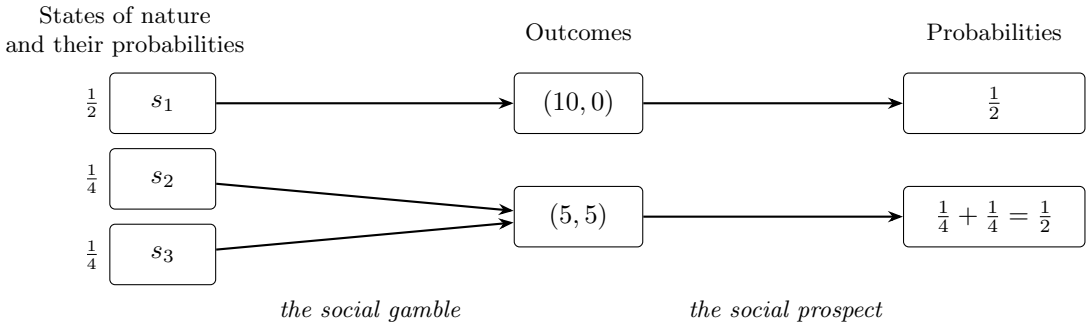



Figure 1: A social gamble and its social prospect. Two states of nature, $s_2$ and $s_3$, result in $(5, 5)$, so the probability of $(5, 5)$ is the sum of theirs.

I will try in these notes to use the correct terminology, but in person I'll probably continue to reflexively say "prospect" when I mean "gamble", because it's easier to say.

## 1 More on Harsanyi

Last week we saw Harsanyi's aggregation theorem: from Anonymity, Statewise Dominance, Ex Ante Pareto, Expected Utility and Transitivity, it derives Utilitarianism, or more strictly VNM utilitarianism (week 7, section 6.9). This section looks at the theorem and its premises more closely.

## 1.1 From Pairwise Anonymity to Anonymity

Anonymity says that rearranging who gets which wellbeing level, in any way at all, gives an equally good outcome. Given Transitivity, it follows from a principle that looks much more modest, since it only concerns swaps between two people.

**Pairwise Anonymity**: If outcome $Y$ can be obtained from outcome $X$ by swapping the wellbeing levels of two people, then $X \sim Y$.

Here is an example with six people, getting from $(1,2,3,4,5,6)$ to $(3,6,1,5,2,4)$ by swapping two levels at a time. The two levels swapped at each step are in bold.

$$\begin{aligned} & (1,2,3,4,5,6) \\ \sim \quad & (\mathbf{3},2,\mathbf{1},4,5,6) \\ \sim \quad & (3,\mathbf{6},1,4,5,\mathbf{2}) \\ \sim \quad & (3,6,1,\mathbf{5},\mathbf{4},2) \\ \sim \quad & (3,6,1,5,\mathbf{2},\mathbf{4}) \end{aligned}$$

Each step is equally good as the one before, by Pairwise Anonymity. So, by Transitivity, $(1,2,3,4,5,6) \sim (3,6,1,5,2,4)$.

The same works for any rearrangement of $n$ people, because any rearrangement can be made by swapping two levels at a time:

(1) swap the level that should end up first into first place;

(2) then swap the level that should end up second into second place;

(3) and so on.

Once the first $n-1$ places are right, the last one must be too, so this takes at most $n-1$ swaps. (In the example, the third place was already right when we got to it, so only four swaps were needed, rather than five.) Each swap gives an equally good outcome, by Pairwise Anonymity, so by Transitivity any rearrangement gives an equally good outcome. That is Anonymity.

This works for any finite number of people. (With infinitely many people, it doesn't, but that need not concern us.)

## 1.2 Two anonymity principles

Write $\pi X$ for the outcome obtained from outcome $X$ by a rearrangement $\pi$ of who gets which wellbeing level. (A rearrangement like this is called a *permutation*.) For example, if $X = (10,0)$ and $\pi$ swaps person 1 and person 2, then $\pi X = (0,10)$. Anonymity can then be written like this:

**Anonymity**: For any rearrangement $\pi$, $\pi X \sim X$.

Here is a second principle, which is weaker (given Transitivity):

**Permutation Invariance**: For any rearrangement $\pi$, $X \succsim Y$ if and only if $\pi X \succsim \pi Y$.

Permutation Invariance rearranges *both* outcomes in the same way, and says that this doesn't change how they compare. For example, Permutation Invariance says that $(10,0) \succsim (5,5)$ if and only if $(0,10) \succsim (5,5)$, since swapping the first two people leaves $(5,5)$ unchanged.

## 1.3 How the two principles differ

Anonymity and Transitivity together give Permutation Invariance. Suppose $X \succsim Y$. By Anonymity, $\pi X \sim X$ and $Y \sim \pi Y$, so we have

$$\pi X \sim X \succsim Y \sim \pi Y,$$

and so, by Transitivity, $\pi X \succsim \pi Y$. For the other direction, suppose $\pi X \succsim \pi Y$. Let $\pi'$ be the rearrangement that undoes $\pi$, putting everyone back where they started, so that $\pi'(\pi X) = X$ and $\pi'(\pi Y) = Y$. The same reasoning, applied to $\pi X$, $\pi Y$ and $\pi'$, gives $X \succsim Y$.

But Permutation Invariance doesn't give Anonymity. To see this, rank outcomes by the Pareto principle alone: $X \succsim Y$ if and only if everyone is at least as well off in $X$ as in $Y$. This ranking satisfies Permutation Invariance, because rearranging both outcomes in the same way doesn't change who is better off where: if everyone is at least as well off in $X$ as in $Y$, the same is true of $\pi X$ and $\pi Y$, and conversely. But under this ranking, $(10, 0)$ and $(0, 10)$ are *incomparable*: neither is at least as good as the other, since person 1 is better off in the first and person 2 is better off in the second. So they are not equally good, and Anonymity fails.

So what does Permutation Invariance rule out? It rules out a rearrangement of an outcome being *better* than the original.

## 1.4 Permutation Invariance: no rearrangement is better

Write $\pi^2 X$ for $\pi(\pi X)$, that is, $\pi$ applied twice; likewise $\pi^3 X$ for $\pi$ applied three times, and so on.

**Claim.** Given Permutation Invariance and Transitivity, $\pi X \succ X$ is impossible.

*Proof.* Suppose, for contradiction, that $\pi X \succ X$. That is, $\pi X \succsim X$, but not $X \succsim \pi X$.

Apply Permutation Invariance, rearranging both $\pi X$ and $X$ by $\pi$. Since $\pi X \succsim X$, we get $\pi^2 X \succsim \pi X$. And since it is not the case that $X \succsim \pi X$, it is not the case that $\pi X \succsim \pi^2 X$ either: Permutation Invariance is an "if and only if", so if $\pi X \succsim \pi^2 X$ held, so would $X \succsim \pi X$. So $\pi^2 X \succ \pi X$. Applying Permutation Invariance again in the same way gives $\pi^3 X \succ \pi^2 X$, and so on.

Now, for some number $k$, $\pi^k X = X$: every rearrangement of finitely many people, repeated enough times, gets everyone back to where they started. (This is basically the cycle decomposition theorem. Informally: follow any one person's level as $\pi$ is applied over and over. Since there are only finitely many places, and $\pi$ never sends two levels to the same place, it eventually comes back to where it started, and it then goes round the same cycle again. Each level has a cycle of some length; repeat $\pi$ a number of times that is a multiple of every one of these lengths, such as their least common multiple, and every level is back where it started.)

So we have a chain

$$X = \pi^k X \succ \pi^{k-1} X \succ \cdots \succ \pi^2 X \succ \pi X \succ X.$$

By Transitivity (which carries over from $\succsim$ to $\succ$), $X \succ X$. That is a contradiction, since nothing is better than itself. So $\pi X \succ X$ is impossible. □

## 1.5 With Completeness: Anonymity

The claim tells us that $\pi X$ is not better than $X$. And $X$ is not better than $\pi X$ either: apply the claim to the outcome $\pi X$ and the rearrangement $\pi'$ that undoes $\pi$. Since $\pi'(\pi X) = X$, the claim says that $X \succ \pi X$ is impossible.

Now suppose we also have Completeness: for any outcomes $X$ and $Y$, either $X \succsim Y$ or $Y \succsim X$. (Completeness was not among last week's five premises.) Then either $X \succsim \pi X$ or $\pi X \succsim X$. If $X \succsim \pi X$, then, since $X$ is not better than $\pi X$, it must also be that $\pi X \succsim X$. Likewise, if $\pi X \succsim X$, then, since $\pi X$ is not better than $X$, it must also be that $X \succsim \pi X$. Either way, both hold, so $X \sim \pi X$. That's Anonymity.

So, given Transitivity and Completeness, Permutation Invariance gives Anonymity, and we saw in section 1.3 that Anonymity gives Permutation Invariance. The two principles differ only where there is incomparability.

## 1.6 Equally good, or incomparable?

So Permutation Invariance allows two verdicts about $(10, 0)$ and $(0, 10)$. They might be equally good, as Anonymity says; or they might be incomparable: neither better, nor worse, nor equally good.

If they are equally good, then

$$(11, 0) \succ (10, 0) \sim (0, 10),$$

where $(11, 0) \succ (10, 0)$ by the Pareto principle. So, by Transitivity, $(11, 0) \succ (0, 10)$: person 1 getting 11 is better than person 2 getting 10. That's not obvious. Incomparability can survive small improvements: $(11, 0)$ might still be neither better nor worse than $(0, 10)$.

## 1.7 Keeping Permutation Invariance, dropping Anonymity

In the simple case of the proof (week 7, section 4.1), Anonymity and Statewise Dominance gave Unequal $\sim$ Equal Chances. Here are the two gambles again (Table 1).

Table 1: The first move of the proof.

| | Heads | Tails |
|---|---|---|
| Unequal | $(10, 0)$ | $(10, 0)$ |
| Equal Chances | $(10, 0)$ | $(0, 10)$ |

If the coin lands Heads, the two gambles give the same outcome. If it lands Tails, Unequal gives $(10, 0)$ and Equal Chances gives $(0, 10)$. If those two outcomes are incomparable, rather than equally good, Statewise Dominance no longer tells us that Unequal and Equal Chances are equally good.

So dropping Anonymity, while keeping Permutation Invariance, blocks the first move of the proof.

But could Equal Chances then be better than Unequal, as the fairness intuition says (week 7, section 5.2)? If it were, it would be better than $(10, 0)$ for sure, even though in each state its outcome is $(10, 0)$ or $(0, 10)$, neither of which is better than $(10, 0)$. That violates the following principle:

**Negative Dominance**: A gamble can't be better than another unless at least some outcome of the first is better than at least some outcome of the second.[1]

Many people find Negative Dominance quite plausible. So perhaps one could avoid the theorem's conclusion by dropping Anonymity. But if the aim is to say, with the fairness intuition, that Equal Chances is better than Unequal, this means giving up Negative Dominance. So it's not so clear that that's a good way of doing it.

## 1.8 Ex Ante Pareto, one person at a time

The second move of the proof, from Equal Chances to Equal (week 7, section 4.1), can be taken in two steps, changing one person's gamble at a time (Table 2).

Table 2: The second move, one person at a time.

| | Heads | Tails |
|---|---|---|
| Equal Chances | $(10, 0)$ | $(0, 10)$ |
| Person 1's gamble changed | $(5, 0)$ | $(5, 10)$ |
| Equal | $(5, 5)$ | $(5, 5)$ |

As in week 7, the numbers are utilities. In Equal Chances, person 1's expected utility is $\frac{1}{2} \cdot 10 + \frac{1}{2} \cdot 0 = 5$. First, give person 1 utility 5 however the coin lands, and leave person 2's gamble as it is. Person 1's expected utility is still 5, so by Expected Utility the change is equally good for person 1; and nothing changes for

[1] The name and the principle are from Lederman, Harvey (forthcoming). "Of Marbles and Matchsticks". In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler et al. Oxford: Oxford University Press. There it is stated for a single person.

person 2. Then do the same for person 2: person 2's expected utility is $\frac{1}{2} \cdot 0 + \frac{1}{2} \cdot 10 = 5$, and giving person 2 utility 5 however the coin lands leaves it at 5. That gives EQUAL. At each step, the new gamble is equally good for everyone as the old one, so by Ex Ante Pareto, applied both ways round, the two are equally good. By Transitivity, EQUAL CHANCES $\sim$ EQUAL. The same result holds more generally.

Each step changed only one person's gamble. So, using Transitivity to chain the two steps together, the proof only needs Ex Ante Pareto for changes to one person at a time: that is, Ex Ante Pareto applied only to pairs of gambles that differ in what they give one person.

## 2 Different numbers of people

So far, every outcome has had the same people in it. What if the number of people can vary? Here is one principle about such cases. We set the zero of the wellbeing scale at the level at which a life would be neither good nor bad, and say that a life is *neutral* if it has wellbeing 0.

**Neutral Addition**: Adding people whose lives are neutral leaves an outcome equally good.

If we add Neutral Addition to the theorem's premises, we get *total utilitarianism*: $X \succsim Y$ if and only if $X$ has at least as great a total of wellbeing as $Y$ (strictly, of utilities: week 7, section 6), whatever the numbers of people in $X$ and $Y$.

Here's why. Take any two outcomes, $X$ and $Y$, and some person who doesn't exist in either of them. By Neutral Addition, adding that person at 0 to each outcome leaves it equally good. Now each outcome has a fixed set of people, so we can use the theorem: concentrating all of the outcome's wellbeing in the added person, and giving everyone else 0, leaves it equally good, since the total is the same. Then remove everyone else, who is now at 0. This is equally good again, by Neutral Addition: adding them back would leave the outcome equally good. So $X$ is equally good as the outcome in which only the added person exists, with the whole of $X$'s total; and likewise for $Y$. These two outcomes contain the very same person, so the one with the greater total is better, by the Pareto principle. This works even if $X$ and $Y$ have nobody in common.

Table 3 does this for two outcomes, $X$, in which persons 1, 2 and 3 exist, and $Y$, in which only persons 1 and 2 exist. The added person is person 4, who exists in neither. $\Omega$ means that the person doesn't exist in that outcome. A person at 0 exists, with a neutral life.

Table 3: Add a neutral life, concentrate, then remove the neutral lives.

| | Outcome | Total | Why |
|---|---|---|---|
| $X$ | $(3, 1, 2, \Omega)$ | 6 | |
| $\sim$ | $(3, 1, 2, 0)$ | 6 | Neutral Addition |
| $\sim$ | $(0, 0, 0, 6)$ | 6 | The theorem |
| $\sim$ | $(\Omega, \Omega, \Omega, 6)$ | 6 | Neutral Addition |
| $Y$ | $(2, 5, \Omega, \Omega)$ | 7 | |
| $\sim$ | $(2, 5, \Omega, 0)$ | 7 | Neutral Addition |
| $\sim$ | $(0, 0, \Omega, 7)$ | 7 | The theorem |
| $\sim$ | $(\Omega, \Omega, \Omega, 7)$ | 7 | Neutral Addition |

Now $(\Omega, \Omega, \Omega, 7) \succ (\Omega, \Omega, \Omega, 6)$ by the Pareto principle, since person 4 at 7 is better off than person 4 at 6. So, by Transitivity,

$$Y \sim (\Omega, \Omega, \Omega, 7) \succ (\Omega, \Omega, \Omega, 6) \sim X,$$

and $Y \succ X$: the greater total wins, even though $X$ has more people.

For more on this, see Tutorial 4 Exercises, Question 1.

# 3 Egalitarianism

## 3.1 Two versions

Parfit (1997) distinguishes two versions of egalitarianism.[2]

**Egalitarianism (deontic)**: Inequality should be avoided.

**Egalitarianism (telic)**: It is in itself bad if some people are worse off than others.

"Telic" is short for "teleological": the telic version is a view about how good outcomes are. The deontic version is a view about what we ought to do. As Parfit puts it, on the telic view inequality is bad; on the deontic view it is unjust (Parfit 1997: 208).

In principle, our social welfare ordering $\succsim$ could be about either. Recall from week 7 (section 3.1) that the social welfare ordering can be read in several ways. On one reading, $X \succsim Y$ says that $X$ is at least as *good* as $Y$; on another, it says something about which outcome *ought to be chosen*. I'll mostly have the betterness interpretation in mind.

## 3.2 Egalitarianism at work

Take the two outcomes from last week, in Table 4.

Table 4: Two outcomes with the same total.

| | Outcome | Total |
|---|---|---|
| Unequal | $(10, 0)$ | 10 |
| Equal | $(5, 5)$ | 10 |

According to Utilitarianism, they are equally good, since they have the same total. According to telic egalitarianism, Unequal is worse: person 2 is worse off than person 1, and that is bad in itself. According to deontic egalitarianism, Unequal should not be chosen.

## 3.3 Equality for its own sake

If we claim that equality is good, we may mean only that it has good effects. Inequality can produce conflict, damage the self-respect of those who are worst off, or put some people in the power of others. If that is all we mean, then equality has *instrumental* value: it is good as a means.

For true egalitarians, equality has *intrinsic* (or *final*) value: it is good in itself.[3] And if inequality is in itself bad, it is bad even when it has no bad effects (Parfit 1997: 206).

## 3.4 The Divided World

Parfit (1997: 206) illustrates this with what he calls the Divided World. Imagine that the world is divided into two halves, and neither half knows about or has any contact with the other. Consider the following possible social outcomes in Table 5.

[2] Parfit, Derek (1997). "Equality and Priority". In: *Ratio* 10.3, pp. 202–221. He calls the telic statement the *Principle of Equality* (p. 204). The optional readings for this week are Parfit (1997); Temkin, Larry S. (2000). "Equality, Priority, and the Levelling Down Objection". In: *The Ideal of Equality*. Ed. by Matthew Clayton and Andrew Williams. London/New York: Macmillan and St Martin's Press. Chap. 6, pp. 126–161; and Adler, Matthew D. and Nils Holtug (2019). "Prioritarianism: A Response to Critics". In: *Politics, Philosophy & Economics* 18.2, pp. 101–144.

[3] Strictly, the opposite of instrumental value is final value, not intrinsic value; the opposite of intrinsic value is extrinsic value. See Korsgaard, Christine M. (1983). "Two Distinctions in Goodness". In: *The Philosophical Review* 92.2, pp. 169–195. I'll keep saying "intrinsic", because it is more convenient, but what I strictly mean is final value.

Table 5: The Divided World.

| | One half | Other half | Average |
|---|---|---|---|
| (1) | 100 | 200 | 150 |
| (2) | 145 | 145 | 145 |

Outcome (1) has more wellbeing. But many people think (2) is better. The two halves are unconnected, so the inequality in (1) has no effects. So if (2) is better because it is more equal, inequality must be bad in itself. That is the telic view, and this kind of case provides one motivation for adopting it rather than the deontic view.

Another motivation is that we would like to say that inequality can be bad even if there's nothing to be done to prevent it. For example, it seems regrettable that some people have lower average lifespans than others. Telic egalitarianism straightforwardly accounts for this, but deontic egalitarianism either cannot or has to tell a more complicated story than "the inequality is still bad".

### 3.5 Levelling down

Consider the change in Table 6.

Table 6: Levelling down.

| | Outcome |
|---|---|
| Before | $(10, 5)$ |
| After | $(5, 5)$ |

Nobody gains, and person 1 loses 5. According to telic egalitarianism, the change is in one way good, since it removes the inequality, even if it is worse all things considered.

The *Levelling Down Objection* says that there is nothing good about the change at all (Parfit 1997: 210–211). It's not just that levelling down would be wrong: the objection is that it isn't good in *any* way. It seems that egalitarians can't straightforwardly endorse that claim.

## 4 Prioritarianism

### 4.1 The Priority View

Here is Parfit's statement of the alternative to egalitarianism.

**The Priority View**: Benefiting people matters more the worse off these people are (Parfit 1997: 213).

For a utilitarian, how much a benefit matters depends only on how big it is. For a prioritarian, it also depends on how well off the person who gets it is. The priority doesn't have to be absolute: a big enough benefit to someone who is better off might outweigh a small benefit to someone who is worse off.

### 4.2 Not a view about equality

Egalitarians care about *relations*: how each person's level compares with other people's. Prioritarians care only about *absolute levels*: how well off each person is.

Parfit's comparison is this:

> People at higher altitudes find it harder to breathe. (Parfit 1997: 214)

That's because of how high up they are, not because other people are lower down: they would find it just as hard to breathe if there were nobody lower down. In the same way, on the Priority View, benefits to the worse off matter more because these people are at a lower absolute level, not because others are better off.

One can be both an egalitarian and a prioritarian at the same time: it's perfectly coherent to think that inequality is bad *and*, even setting inequality aside, benefiting the worse off matters more. But a prioritarian, as such, doesn't think that inequality is bad in itself.

### 4.3 Two reasons to prefer EQUAL

In Table 4, the egalitarian and the prioritarian can both prefer EQUAL to UNEQUAL, but for different reasons.

According to the egalitarian, in UNEQUAL, person 2 is worse off *than person 1*, and that is bad in itself. According to the prioritarian, in UNEQUAL, person 2 is *badly off*, so benefits to person 2 matter more, whatever person 1 has. Prioritarians favour equality only as a by-product of giving priority to the worse off.

A utilitarian can also prefer equal *resources*, since extra resources usually do less for someone who already has a lot. But a utilitarian doesn't prefer equal *wellbeing*: according to Utilitarianism, UNEQUAL and EQUAL are equally good.

### 4.4 Priority and levelling down

In the levelling down case of Table 6, person 1 goes from 10 to 5, and person 2 stays at 5. According to prioritarianism, that is worse, and in no way better, since nobody's absolute level goes up. So prioritarianism avoids the Levelling Down Objection (Parfit 1997: 214).

And it can still say that (2) is better than (1) in the Divided World (Table 5). In the change from (1) to (2), the better off would lose more than the worse off would gain, but the worse off's gains count for more (Parfit 1997: 221).

### 4.5 Separability over persons

Here is a principle that marks a structural difference between the two views, from Adler and Holtug (2019: 104), who call it Person Separability.

**Separability over Persons**: How two outcomes compare doesn't depend on the wellbeing of people who are equally well off in both.

We can write it this way, using variables $X$, $X'$, $Y$ and $Y'$ to denote a population or wellbeing distribution, and the $+$ symbol to denote conjoining two wellbeing distributions together, so that $X + Y$ consists of the $X$-people at their levels in $X$ together with the $Y$-people at their levels in $Y$:

**Separability over Persons**: For all wellbeing distributions $X, X'$, $Y$ and $Y'$, if $X$ and $X'$ contain the same people, $Y$ and $Y'$ contain the same people, and nobody is among both the $X$-people and the $Y$-people, then

$$X + Y \succsim X' + Y \iff X + Y' \succsim X' + Y'$$

In other words, it doesn't matter whether the unaffected people are at the $Y$ levels or at the $Y'$ levels. If those wellbeing levels don't change when we're comparing $X$ to $X'$, we can ignore them in evaluating the overall outcome.

Prioritarianism has this built in. How much a benefit to someone matters depends only on how well off that person is, whatever anyone else has. In particular, it makes no difference whether some other unaffected people are at their $Y$ or their $Y'$ levels.

Egalitarianism needn't satisfy Separability over Persons. Whether a change reduces inequality can depend on where the unaffected people are. Some egalitarian views happen to satisfy it; others don't.[4]

[4] Adler and Holtug themselves count a view as egalitarian only if it violates Separability over Persons (p. 104).

By the way, remember the analogy between persons and states of nature? If we switch around Separability over Persons and make it about states of nature instead, we would get "Separability over States", which turns out to be the same thing as the sure-thing principle (week 4, section 6.2). (When evaluating gambles, we don't need to pay attention to what happens in any states where the outcome doesn't depend on which of the two gambles we pick.)

## 4.6 Separability over persons: an example

Table 7 is based on an example of Adler and Holtug (2019: 106).

Table 7: Person 3 is unaffected.

| | Outcome |
|---|---|
| $X$ | $(4, 36, 49)$ |
| $Y$ | $(9, 16, 49)$ |
| $X'$ | $(4, 36, 1)$ |
| $Y'$ | $(9, 16, 1)$ |

Both times, going from $X$ to $Y$ and from $X'$ to $Y'$, person 1 gains 5 and person 2 loses 20. The only difference is person 3's level, and person 3 is unaffected. So Separability over Persons says that $X$ and $Y$ must compare in the same way as $X'$ and $Y'$.

A prioritarian satisfies this. Prioritarianism, written down precisely in section 4.8, adds up a term for each person that depends only on that person's own wellbeing, so person 3's term is the same in $X$ and $Y$ and cancels out. For example, if we use the square root priority weighting function, then persons 1 and 2 contribute $\sqrt{4} + \sqrt{36} = 2 + 6 = 8$ in $X$ and $X'$, against $\sqrt{9} + \sqrt{16} = 3 + 4 = 7$ in $Y$ and $Y'$. So $X$ beats $Y$, and $X'$ beats $Y'$.

But an egalitarian view could rank $Y$ above $X$ but $X'$ above $Y'$. On such a view, person 1's gain carries extra weight when person 1 is the worst off, as in $X$; but in $X'$ the worst off is person 3, and person 1's gain no longer carries that extra weight.

## 4.7 A representation theorem for prioritarianism

Adler and Holtug (2019: 103–104) state a representation theorem for prioritarianism. It assumes a fixed population. Suppose that a complete and transitive ranking of outcomes satisfies these conditions:

- Strong Pareto: if an outcome is better for someone and at least as good for everyone, it is better;
- Pareto Indifference: if everyone is equally well off in two outcomes, they are equally good;
- Anonymity;
- Pigou–Dalton: a pure transfer of wellbeing from someone better off to someone worse off (the same amount lost by one and gained by the other), which doesn't swap their ranks and leaves everyone else unaffected, is an improvement;
- Separability over Persons;
- Continuity: small enough changes to an outcome don't reverse a strict ranking.

Then it ranks outcomes by $f(w_1) + \cdots + f(w_n)$, for some strictly increasing, strictly concave function $f$. And every ranking of that form satisfies all of the conditions.

**WARNING**: The Pareto principles here are about comparing social outcomes, not about social gambles. Ex Ante Pareto is much stronger and correspondingly more controversial.

### 4.8 Writing prioritarianism down

With $n$ people, numbered 1 to $n$, and $w_i(X)$ person $i$'s wellbeing in outcome $X$:

**Prioritarianism**:

$$X \succsim Y \quad \text{iff} \quad f\big(w_1(X)\big) + \cdots + f\big(w_n(X)\big) \;\geq\; f\big(w_1(Y)\big) + \cdots + f\big(w_n(Y)\big),$$

where $f$ is increasing and strictly concave.

That is, $X \succsim Y$ if and only if $\sum_i f\big(w_i(X)\big) \geq \sum_i f\big(w_i(Y)\big)$. We generally call $f$ the *priority weighting function*, and we can call $\sum_i f\big(w_i(X)\big)$ the *priority-weighted total*.

The priority weighting function being strictly concave means that each extra unit of wellbeing adds less to the priority-weighted total than the one before. So a benefit counts for more, overall, when it goes to someone who is worse off. The priority weighting function is just a way of weighting benefits to the worse off more heavily than comparable benefits to the better off.

For example, take $f(w) = \sqrt{w}$. This is not the true one, just a convenient example for non-negative wellbeing. Then:

– $(25, 25)$ has priority-weighted total $\sqrt{25} + \sqrt{25} = 5 + 5 = 10$;

– $(49, 1)$ has priority-weighted total $\sqrt{49} + \sqrt{1} = 7 + 1 = 8$.

So $(25, 25)$ is better than $(49, 1)$, though the totals are the same.

### 4.9 Prioritarianism and Utilitarianism can disagree

Prioritarianism doesn't only break ties between outcomes with the same total. Consider Table 8.

Table 8: Totals and priority-weighted totals.

| | Outcome | Total | Priority-weighted total ($f(w) = \sqrt{w}$) |
|---|---|---|---|
| $X$ | $(25, 0)$ | 25 | $5 + 0 = 5$ |
| $Y$ | $(9, 9)$ | 18 | $3 + 3 = 6$ |

According to Utilitarianism, $X$ is better, since it has the greater total. According to prioritarianism with $f(w) = \sqrt{w}$, $Y$ is better. Going from $Y$ to $X$, person 1 gains 16 and person 2 loses 9. Person 2's loss counts for more, because person 2 ends up so badly off.

## 5 Ex ante and ex post

### 5.1 Ex ante and ex post prioritarianism

So far, prioritarianism ranks outcomes. To rank gambles, we have to decide how to apply the priority when there is risk. There are two natural ways to do it (there are others). In this section, each person's betterness is risk neutral about wellbeing: utility is wellbeing.

**Ex Ante Prioritarianism**: Work out each person's expected wellbeing. Then take the sum of $f$ of those.

**Ex Post Prioritarianism**: Work out the priority-weighted total, the sum of $f$(wellbeing), in each state of nature. Then take the expectation.

Ex post prioritarianism is the more common of the two. To see how they work, we'll use three gambles: $(49, 1)$ for sure; a coin toss giving $(49, 1)$ if Heads and $(1, 49)$ if Tails; and $(25, 25)$ for sure. As before, $f(w) = \sqrt{w}$.

## 5.2 Ex ante prioritarianism at work

Table 9: The coin toss.

| Heads | Tails |
|---|---|
| $(49, 1)$ | $(1, 49)$ |

In the coin toss (Table 9), each person's expected wellbeing is $\frac{1}{2} \cdot 49 + \frac{1}{2} \cdot 1 = 25$. The sum of $f$ of those is $\sqrt{25} + \sqrt{25} = 5 + 5 = 10$. $(49, 1)$ for sure scores $\sqrt{49} + \sqrt{1} = 7 + 1 = 8$. So, according to ex ante prioritarianism, the coin toss is better than $(49, 1)$ for sure.

But the coin toss is no better than $(49, 1)$ for sure in either state of nature. If Heads, the outcome is the same. If Tails, it is $(1, 49)$, which is equally good as $(49, 1)$ according to prioritarianism, since it has the same priority-weighted total. So ex ante prioritarianism violates Statewise Dominance.

## 5.3 Ex post prioritarianism at work

According to ex post prioritarianism, the coin toss scores $\sqrt{49} + \sqrt{1} = 7 + 1 = 8$ if Heads, and $\sqrt{1} + \sqrt{49} = 8$ if Tails. So its expectation is 8. $(25, 25)$ for sure scores $5 + 5 = 10$ in every state, so its expectation is 10. So $(25, 25)$ for sure is better than the coin toss.

But each person's expected wellbeing is 25 either way, and since each person's betterness is risk neutral about wellbeing, the two gambles are equally good for each person. So ex post prioritarianism violates Ex Ante Pareto.

## 5.4 The two versions at work

Table 10 puts the scores side by side.

Table 10: Scores with $f(w) = \sqrt{w}$.

| | $(49, 1)$ for sure | Coin toss | $(25, 25)$ for sure |
|---|---|---|---|
| Ex ante | 8 | 10 | 10 |
| Ex post | 8 | 8 | 10 |

The ex ante version goes against Statewise Dominance, and the ex post version goes against Ex Ante Pareto. So each version blocks a different move in the proof.

## 5.5 Ex post prioritarianism with one person

Ex post prioritarianism violates Ex Ante Pareto even when there is only one person. Suppose only one person exists, and her betterness is risk neutral about wellbeing. Compare the two gambles in Table 11.

Table 11: One person.

| | Heads | Tails |
|---|---|---|
| Sure thing | 25 | 25 |
| Gamble | 49 | 1 |

Her expected wellbeing is 25 either way, so the two gambles are equally good for her. According to ex post prioritarianism, the sure thing scores $\sqrt{25} = 5$, and the gamble scores $\frac{1}{2} \cdot 7 + \frac{1}{2} \cdot 1 = 4$. So the sure thing is better. That violates Ex Ante Pareto, with nobody else around: there is no inequality and no unfairness.[5]

[5]Greaves, Hilary (2015). "Antiprioritarianism". In: *Utilitas* 27.1, pp. 1–42. The point is on p. 11.

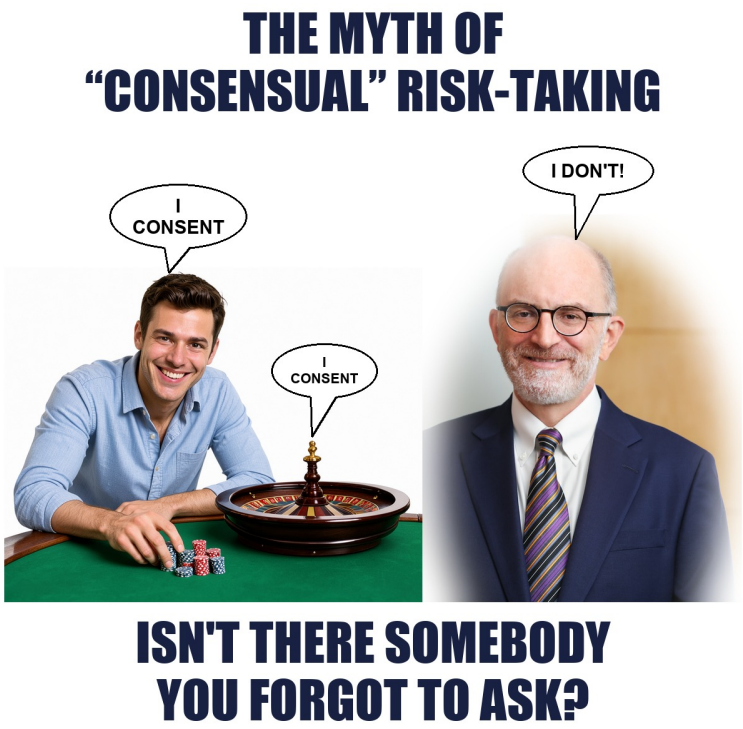



Figure 2: Professor Matthew Adler (Duke), a prominent prioritarian, applying the morally correct degree of risk aversion in a single-person case.

### 5.6 Ex ante and ex post inequality

There are two kinds of inequality to distinguish when there is risk. There is *ex post* inequality when, in some state of nature, people end up with different wellbeing. There is *ex ante* inequality when people face gambles with different expected wellbeing.

Table 12 goes back to the three gambles from last week.

Table 12: Ex ante and ex post equality.

| | Heads | Tails | Ex ante equal? | Ex post equal? |
|---|---|---|---|---|
| Unequal | $(10, 0)$ | $(10, 0)$ | No | No |
| Equal Chances | $(10, 0)$ | $(0, 10)$ | Yes | No |
| Equal | $(5, 5)$ | $(5, 5)$ | Yes | Yes |

In Equal Chances, each person's expected wellbeing is 5, so there is ex ante equality. But whichever way the coin lands, one of them gets 10 and the other gets 0, so there is ex post inequality.

### 5.7 Ex ante equality, any ex post inequality

Take any outcome, say $(1, 2, 3)$. Spread it over three equally likely states, rotating who gets what, as in Table 13. This is a lottery-equalisation of $(1, 2, 3)$, in the sense of last week (week 7, section 4.2).

Table 13: A lottery-equalisation of $(1, 2, 3)$. Each state has probability $\frac{1}{3}$.

| State 1 | State 2 | State 3 |
|---|---|---|
| $(2, 3, 1)$ | $(3, 1, 2)$ | $(1, 2, 3)$ |

Each person gets 1, 2 and 3, once each, so everyone's expected wellbeing is $\frac{1}{3} \cdot 1 + \frac{1}{3} \cdot 2 + \frac{1}{3} \cdot 3 = 2$. That's ex ante equality. But in every state, one person has 1, one has 2 and one has 3, so every state is as unequal as $(1, 2, 3)$. And this works for any outcome, however unequal. So ex ante equality is compatible with any amount of ex post inequality.

### 5.8 Ex post equality guarantees ex ante equality

Now suppose every state is perfectly equal: in each state, everyone has the same wellbeing as everyone else. Then everyone faces the very same gamble, so everyone has the same expected wellbeing. That's ex ante equality.

It doesn't go the other way round: Equal Chances is ex ante equal, but unequal in every state. So, with a fixed population, guaranteed ex post equality is the more demanding kind.

### 5.9 Ex ante and ex post egalitarianism

Egalitarians face the same choice as prioritarians: should they care about inequality in people's chances, or inequality in outcomes?

According to ex ante egalitarianism, Equal Chances gives both people the same chances (each has a 50–50 chance of 10 or 0), so it is better than Unequal. That's Diamond's judgement, which goes against Statewise Dominance (week 7, section 5.2).

According to ex post egalitarianism, Equal Chances guarantees an unequal outcome, so it is worse than Equal. That's the objection to Ex Ante Pareto (week 7, section 5.3).

These aren't exclusive. An egalitarian can care about both ex ante and ex post inequality; likewise, a prioritarian can combine the ex ante and ex post versions.

## 6 Priority and risk aversion

### 6.1 Risk-averse betterness

Recall from week 7 (sections 6.3 and 6.4) that utility might be a curved function of wellbeing. If it is concave, each person's betterness relation is risk averse about wellbeing; if it is convex, it is risk seeking. In section 5 we assumed that utility is wellbeing. Now suppose instead that everyone's utility is $\sqrt{\text{wellbeing}}$, as in Table 14.

Table 14: Everyone's utility is $\sqrt{\text{wellbeing}}$.

| Wellbeing | 0 | 4 | 16 |
|---|---|---|---|
| Utility | 0 | 2 | 4 |

A 50–50 chance of 16 or 0 has expected utility $\frac{1}{2} \cdot 4 + \frac{1}{2} \cdot 0 = 2$. That's the utility of 4 for sure. So, for each person, the gamble is exactly as good as 4 for sure.

### 6.2 The proof again, with risk-averse betterness

Now run the simple case of the proof again, with a fair coin and these utilities, starting from $(16, 0)$ for sure (Table 15).

Table 15: The proof again. Everyone's utility is $\sqrt{\text{wellbeing}}$.

| | Heads | Tails | Why |
|---|---|---|---|
| $(16, 0)$ for sure | $(16, 0)$ | $(16, 0)$ | |
| $\sim$ coin toss | $(16, 0)$ | $(0, 16)$ | Anonymity, Statewise Dominance |
| $\sim (4, 4)$ for sure | $(4, 4)$ | $(4, 4)$ | Expected Utility, Ex Ante Pareto |

The first move is unchanged: by Anonymity, $(0, 16)$ is equally good as $(16, 0)$, so by Statewise Dominance the coin toss is equally good as $(16, 0)$ for sure. The second move is where the risk aversion comes in. In the coin toss, each person has a 50–50 chance of 16 or 0, and that is as good for them as 4 for sure. So, by Expected Utility and Ex Ante Pareto, the coin toss is equally good as $(4, 4)$ for sure. By Transitivity, $(16, 0) \sim (4, 4)$.

## 6.3 The same steps, with prioritarianism

Now let everyone's utility be wellbeing again, and use ex post prioritarianism with $f(w) = \sqrt{w}$. Table 16 works out the priority-weighted total of the same three gambles in each state.

Table 16: Ex post prioritarianism with $f(w) = \sqrt{w}$. Everyone's utility is wellbeing.

| | Heads | Tails | Expectation |
|---|---|---|---|
| $(16, 0)$ for sure | $4 + 0 = 4$ | $4 + 0 = 4$ | 4 |
| coin toss | $4 + 0 = 4$ | $0 + 4 = 4$ | 4 |
| $(4, 4)$ for sure | $2 + 2 = 4$ | $2 + 2 = 4$ | 4 |

All three are equally good: exactly the verdicts of the proof in Table 15. In Table 15, the square root discounts a 50–50 chance of 16 or 0 to 4 for sure; here it discounts $(16, 0)$ to $(4, 4)$. Risk costs a person exactly what inequality costs overall.

## 6.4 Priority equals risk aversion

In the theorem, the individual utilities do two jobs. Their *expectation* says what is better for each person. Their *sum* says what is better overall.

Suppose everyone's utility is the same strictly concave function of wellbeing: utility $= f(\text{wellbeing})$. Then the theorem's conclusion, that outcomes are ranked by their total utility, says that outcomes are ranked by $f(w_1) + \cdots + f(w_n)$. That is prioritarianism with that $f$.

What about gambles? Take any gamble. By Step 3 of the proof (week 7, section 4.2), the outcome in each state of nature is equally good as its equalisation, in which everyone gets the average utility of that outcome. So, by Statewise Dominance, the gamble is equally good as the gamble that gives each state's equalisation in that state. In that gamble, everyone faces the same gamble over utilities, so by Expected Utility and Ex Ante Pareto it is equally good as the outcome in which everyone gets that gamble's expected utility for sure. That expected utility is the expectation of the outcome's average utility. So comparing two gambles comes down to comparing two outcomes in which everyone is equally well off; and, as in Step 4 of the proof, the one with the higher level is better. So gambles are ranked by the expectation of the average utility, which ranks them in the same way as the expectation of the total utility. That is, the theorem ranks gambles by the expectation of $f(w_1) + \cdots + f(w_n)$: ex post prioritarianism with that $f$.

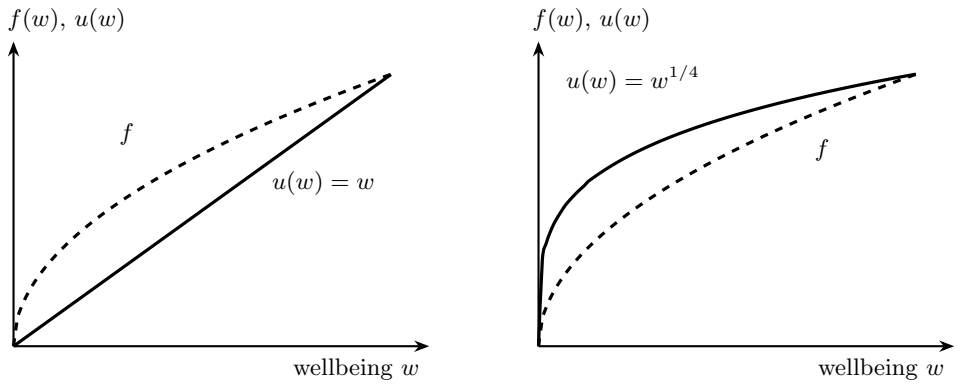
So, given the premises, the priority given to the worse off must be exactly the degree of risk aversion about wellbeing.[6]

[6]Broome (1991: 210–213) makes this point, and calls the resulting view "communal egalitarianism". Broome, John (1991). *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Basil Blackwell.

(From here on, to keep the two roles apart, I'll write $u$ for the utility function and keep $f$ for the prioritarian's priority weighting.)

## 6.5 When prioritarianism disagrees with the theorem

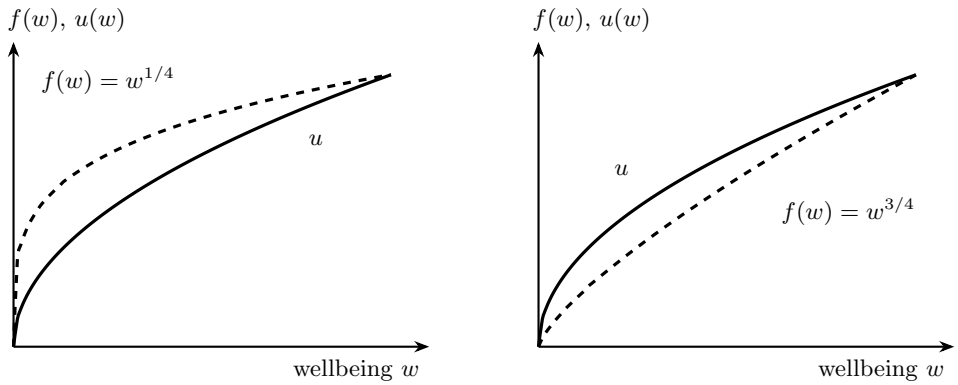
Say everyone's utility is $u$(wellbeing). Then the theorem gives ex post prioritarianism with $f = u$ (or Utilitarianism, if $u$ is a straight line). A prioritarian whose $f$ isn't just a stretched or shifted $u$ disagrees with the theorem, and so has to reject one of its premises.[7] For example, the prioritarian's $f$ could be more concave than $u$: more priority than risk aversion. Or it could be less concave than $u$: less priority than risk aversion, though still priority to the worse off in wellbeing. Figures 3 and 4 show examples.




$f$ more concave than $u$:
more priority than risk aversion

$f$ less concave than $u$:
less priority than risk aversion

Figure 3: Same priority weighting $f(w) = \sqrt{w}$ (dashed), different utilities $u$ (solid).




$f$ more concave than $u$:
more priority than risk aversion

$f$ less concave than $u$:
less priority than risk aversion

Figure 4: Same utility $u(w) = \sqrt{w}$ (solid), different priority weightings $f$ (dashed).

## 6.6 Prioritarianism in terms of utilities

We can also write the prioritarian's verdicts in terms of utilities. Write $v$ for a level of utility: $v = u(w)$. Then wellbeing is $w = u^{-1}(v)$, where $u^{-1}$ is the inverse of $u$: the function that takes each utility back to the wellbeing level that has it. So

$$f(w) = f\big(u^{-1}(v)\big) = (f \circ u^{-1})(v),$$

[7] Compare Greaves (2015: 8–10, 16–17).

where $f \circ u^{-1}$ is the function that applies $u^{-1}$ first and then $f$. The prioritarian's priority-weighted total, written in the people's utilities $v_1, \ldots, v_n$, is then

$$(f \circ u^{-1})(v_1) + \cdots + (f \circ u^{-1})(v_n).$$

That is what it means for $f$ to be more concave than $u$: $f \circ u^{-1}$ is strictly concave.

For example, let $u(w) = \sqrt{w}$, so that $w = v^2$.

– If $f(w) = w^{1/4}$, then $(f \circ u^{-1})(v) = (v^2)^{1/4} = \sqrt{v}$, which is concave.

– If $f(w) = w^{3/4}$, then $(f \circ u^{-1})(v) = (v^2)^{3/4} = v^{3/2}$, which is convex.

In the second case, the prioritarian gives priority to the worse off in wellbeing, but sums a convex function of the utilities: in terms of utilities, the better off get priority. So "priority to the worse off" means different things when it is applied to wellbeing and when it is applied to utilities.

### 6.7 Egalitarianism and the theorem

Whatever the risk attitude, the theorem's conclusion is a sum, with one term for each person. So it satisfies Separability over Persons: an unaffected person's term is the same in both outcomes, and cancels out. So an egalitarian view that isn't separable over persons has to reject a premise of the theorem.

## References


Adler, Matthew D. and Nils Holtug (2019). "Prioritarianism: A Response to Critics". In: *Politics, Philosophy & Economics* 18.2, pp. 101–144.

Broome, John (1991). *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Basil Blackwell.

Greaves, Hilary (2015). "Antiprioritarianism". In: *Utilitas* 27.1, pp. 1–42.

Korsgaard, Christine M. (1983). "Two Distinctions in Goodness". In: *The Philosophical Review* 92.2, pp. 169–195.

Lederman, Harvey (forthcoming). "Of Marbles and Matchsticks". In: *Oxford Studies in Epistemology*. Ed. by Tamar Szabó Gendler et al. Oxford: Oxford University Press.

Parfit, Derek (1997). "Equality and Priority". In: *Ratio* 10.3, pp. 202–221.

Temkin, Larry S. (2000). "Equality, Priority, and the Levelling Down Objection". In: *The Ideal of Equality*. Ed. by Matthew Clayton and Andrew Williams. London/New York: Macmillan and St Martin's Press. Chap. 6, pp. 126–161.