



ARTICLE

# Relative priority

Lara Buchak 

Princeton University, Princeton, NJ, USA  
Email: [lbuchak@princeton.edu](mailto:lbuchak@princeton.edu)

(Received 6 July 2020; revised 6 January 2022; accepted 13 January 2022; first published online 17 June 2022)

## Abstract

The good of those who are worse off matters more to the overall good than the good of those who are better off does. But being worse off than one's fellows is not itself bad; nor is inequality itself bad; nor do differences in well-being matter more when well-being is lower in an absolute sense. Instead, the good of the *relatively* worse-off *weighs more heavily* in the overall good than the good of the relatively better-off does, in virtue of the fact that the former are relatively worse off. This paper articulates and defends the view just described.

**Keywords:** inequality; utilitarianism; egalitarianism; prioritarianism; aggregation

## 1. Introduction

The good of those who are worse off matters more to the overall good than the good of those who are better off does. But being worse off than one's fellows is not itself bad; nor is inequality itself bad; nor do differences in well-being matter more when well-being is lower in an absolute sense. Instead, the good of the *relatively* worse-off *weighs more heavily* in the overall good than the good of the relatively better-off does, in virtue of the fact that the former are relatively worse off. At least, that is my view. The goal of this paper is to articulate this view more precisely and offer a defence.

There are three standard answers in the literature to the question of how to measure the overall good in a society: utilitarianism, egalitarianism and prioritarianism. Utilitarianism says that overall good is total or average good, regardless of how that good is distributed across individuals. Egalitarianism says that overall good is a matter of both total or average good and *also* how that good is distributed, since inequality is in itself bad. Prioritarianism says that overall good is total or average moral value, where additional good accruing to an individual is less morally valuable the better things already are for her. Thus it says that equal distributions are better than unequal ones, but not because inequality is itself bad.

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivatives licence (<http://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial re-use, distribution, and reproduction in any medium, provided that no alterations are made and the original article is properly cited. The written permission of Cambridge University Press must be obtained prior to any commercial use and/or adaptation of the article.

This paper is primarily aimed at those who are motivated by the commitments that motivate prioritarrians – that equal distributions are better than unequal ones, but not because inequality is itself bad. I will show that there is another alternative to prioritarianism, which I call *relative prioritarianism*, and I will give reasons for preferring this alternative.

The positions in this paper each have a philosophical and formal component. However, since I am most concerned with the philosophical motivations that distinguish the positions, I will only briefly introduce the formalisms by way of examples, and leave complete formal definitions to the Appendix.

## 2. Set up and background

We are here concerned with the *aggregation question*: once we've settled on what constitutes an individual's good, how do we evaluate a *distribution* of individual good?

We will use as an example a society with three people. A *social distribution* is a list of the *consequences* for each individual: for example, {ANN, moderately happy life; BOB, short and difficult life; CECIL, extremely happy life} is a distribution (call this distribution D1). (Consequences are to be understood broadly, to include anything that affects individual well-being; for example, they might include an individual's personal relationships.) We want to know how good this distribution is, and whether it is better or worse than other distributions. We will use 'utility' to stand in for *prudential value*, that is, the good of a consequence for the individual who has it, and we will assume that utility is given, whether it tracks some objective fact or is derived from preferences. So, for example, D1 might have utility values {ANN, 200; BOB, 100; CECIL, 300}. To explain the various views, let us also consider another distribution, D2 = {ANN, 200; BOB, 200; CECIL, 200}.

An *aggregation rule* assigns a numerical value to each distribution, such that of two distributions, the one with the greater numerical value is better. Equivalently, an aggregation rule provides a complete ranking of distributions. Whether the ranking or the aggregation rule is more 'fundamental' will not concern us here: we will talk in terms of aggregation rules but keep in mind that rules that give equivalent rankings are in a formal sense the same rule.

We will make some simplifying assumptions. First, we will only consider social distributions with fixed utility: we will not evaluate lotteries over social distributions, or social distributions over lotteries. While any complete theory of aggregation will have to evaluate the latter,<sup>1</sup> I take evaluation of the former to be more fundamental. The second assumption is that we are dealing with a fixed population of individuals. Thus, a given rule can be formulated either as evaluating *total* good or *average* good, since the two formulations will produce equivalent rankings and equivalent values up to scale. We will leave open how the correct rule should be extended to societies with different numbers of people.<sup>2</sup>

<sup>1</sup>See for example the discussion of *ex ante* and *ex post* prioritarianism in McCarthy (2006, 2008). For a survey of *ex ante* and *ex post* choice rules, see d'Aspremont and Gevers (2002: sec. 5).

<sup>2</sup>Since all the rules discussed here are *anonymous* in the sense that they make no reference to the identities of people, they yield the same results for a population of the same size but composed of different people.

The aggregation question presupposes that the good of a distribution can be determined from the good of its constituent parts. This rules out a number of plausible views,<sup>3</sup> but I simply do not have space to address them here.

To answer the aggregation question, we must do two things. We must give a *formal* theory of aggregation, that is, we must provide an aggregation rule. And we must give an *ethical interpretation* of the rule that explains what the rule means in terms of what we value. It is important to remember that formal theories don't come interpreted, and as we will see, it is possible for a formal theory to have more than one possible ethical interpretation, or an ethical claim to have more than one possible formal correlate.

The three most prominent views are utilitarianism, egalitarianism and prioritarianism. Each can be presented in a 'total' or 'average' version.<sup>4</sup>

*Total utilitarianism* says that the value of a distribution is its total utility, and *average utilitarianism* says that the value of a distribution is its average utility. So, the value of both D1 and D2 according to total utilitarianism is 600, and the value of both D1 and D2 according to average utilitarianism is 200. According to utilitarianism, they are equally good.

*Egalitarianism* holds that the fact that some people are worse off than others is bad in itself, independent of the utility values involved. As Parfit defines it, egalitarians accept the Principle of Equality:<sup>5</sup>

Principle of Equality: It is in itself bad if some people are worse off than others.

And they typically accept this principle alongside the Principle of Utility:

Principle of Utility: It is in itself better if people are better off (on average or in total).

Thus, egalitarians hold that utility and equality each matter, though there may be some trade-off involved in doing better with respect to each of these principles.

Formally, egalitarianism takes the average or total (i.e. utilitarian) value of a distribution, and subtracts from it some value that measures the inequality in the distribution.<sup>6</sup> Thus, as long as D1 scores worse than D2 with respect to

<sup>3</sup>Prominent views that this rules out include: views that reject transitivity (e.g. Temkin 1993, 1996), and views according to which the goodness of an option is menu-dependent such as (one interpretation of) the competing claims view (Otsuka and Voorhoeve 2009; Voorhoeve and Fleurbaey 2012; Otsuka 2012; Voorhoeve 2014).

<sup>4</sup>See Appendix 1 for full mathematical details of all the rules in this section.

<sup>5</sup>Parfit (1991: 84). Parfit uses this to define what he calls *Telic* egalitarianism. Another kind of egalitarianism he calls *Deontic* egalitarianism: deontic egalitarians holds that inequality is unjust, i.e. that (when inequality is bad) inequality necessarily involves wrong-doing. Since we are only concerned with evaluations of states of affairs in terms of their goodness, not in evaluating potential paths to these states of affairs, deontic egalitarianism does not concern us here. See also Temkin (2003a, 2003b) on 'noninstrumental' egalitarianism, Holtug (2010: 174), and Hirose (2015: 64-85).

<sup>6</sup>There may be other ways to implement egalitarianism formally (for example, multiply the average or total by some measure of inequality). However, the above formalism is the class most commonly used in philosophical discussion; and since I am here interested in the philosophical position that (in)equality is intrinsically (dis)valuable, discussion won't be changed by using a different class.

inequality, D2 will be better than D1. (For example, taking the average version of the view, if the inequality score of D1 is 30 and that of D2 is 0, then the value of D1 is 170 and the value of D2 is 200.) Egalitarianism says that the overall good in a society is a combination of two different sources of value: prudential value for individuals, and how that value is distributed.

*Prioritarianism* holds that we ought to maximize total *moral value*, where moral value is a concave function of *utility*: it is more morally valuable to increase someone's good by a given amount the lower his absolute level of good.<sup>7</sup> (As one has more utility, additional bits of utility are less morally valuable. As Parfit (1991: 105) says, "just as *resources* have diminishing marginal *utility*, so *utility* has diminishing marginal *moral importance*.") The value of a distribution is then its total moral value. Although the 'total' formulation is normally used, the 'average' formulation is easily defined: the value of a distribution is its average moral value.

So, for example, we might have the following moral value function of utility, which diminishes marginally:

$$v(100) = 100, v(200) = 190, v(300) = 250$$

The overall moral value, according to the total prioritarian, of D1 and D2 will be:

$$P_{total}(D1) = v(200) + v(100) + v(300) = 540$$

$$P_{total}(D2) = v(200) + v(200) + v(200) = 570$$

And the overall moral value of D1 and D2, according to the average prioritarian, will be:

$$P_{average}(D1) = (1/3)v(200) + (1/3)v(100) + (1/3)v(300) = 180$$

$$P_{average}(D2) = (1/3)v(200) + (1/3)v(200) + (1/3)v(200) = 190$$

D2 is better than D1, because D2 contains more total or average moral value: when a distribution is more spread out in terms of utility, individuals with lower utility bring down the total or average *moral value* more than individuals with higher utility bring it up. Thus, if moral value diminishes marginally in utility, then when two distributions have the same total or average utility, the one that is less spread out in terms of utility will have a higher total or average moral value. (Equivalently: it is always better to give a fixed utility benefit to a worse-off person than to a better-off person.)

<sup>7</sup>See Parfit (1991, 2012). See also Broome (1991), Nagel (1991), Arneson (2000), Scheffler (2000), Rabinowicz (2002), Jensen (2003), McCarthy (2006, 2008), Otsuka and Voorhoeve (2009), Adler (2012, 2019), Hirose (2015) and Otsuka (2015). Tungodden (2003) considers variants of prioritarianism, but takes it to be a core principle that prioritarianism adheres to strong separability, a principle I will discuss later. McCarthy (2013) cites the above definition as the "risk-free" definition, though goes on to argue that this definition doesn't ultimately make sense.

### 3. Prioritarian desiderata

Prioritarians accept three claims. The first is about what utility measures, and is accepted by all three parties in the debate so far:

INDIVIDUAL GOOD: Utility for each person depends only on her own consequence.

For example, if the consequence I have has a particular utility for me, then its utility doesn't go down if someone else has a higher level of utility.

INDIVIDUAL GOOD says that *individual utility values* are not relative: whatever is meant by prudential good or well-being, it is not the kind of thing that I get less of because someone else has more of it. This isn't to say that what others have can't make a difference to my well-being. It is that these facts are *already* included in the description of my consequence, the thing to which utility attaches.<sup>8</sup> (So, for example, if the fact that others have much more than I do means I am powerless and this makes my life less good, then the fact that I am powerless is included in the description of the consequence I get.) This may mean that consequence descriptions are somewhat complex; but once we've determined my utility, it doesn't make a difference to me if other people have more or less of *that*. One could think of this either as a substantive point or a stipulative definition of utility.<sup>9</sup>

The second claim that prioritarians accept is the one that egalitarians accept and utilitarians reject:

SPREAD AVERSION: Of two distributions with the same total or average utility value, the one that is less spread-out in terms of individuals' utility levels is better.<sup>10</sup>

To avoid confusion, we should contrast the concept of utility-level spread with the concept of 'welfare diffusion' (Persson 2011, 2012). Welfare diffusion means that a given amount of utility is divided among a larger number of people (e.g. {A, 200; B, 200} is more diffuse than {A, 400}). Utility-level spread means roughly that the individuals' utility levels are 'farther apart' (e.g. {A, 400, B, 0} is more spread out than A, 200; B, 200}).

<sup>8</sup>See e.g. Broome (1991).

<sup>9</sup>A brief remark on where this locates *relational egalitarians* in this debate. Relational egalitarians are philosophers who hold that what is bad about inequality is that it harms people's ability to relate as equals. On the way I've carved up the space, relational egalitarians can count as accepting this principle in a 'stipulative' way, because utility will be defined as an individual's good taking into account her relationships with others. This makes the issue that relational egalitarians are concerned with one of how to determine an individual's good given a complete description of her situation, including what others have (an extremely important issue, but not the focus of this paper). Relational egalitarians still have to answer the main question posed by this paper – once we know how good things are for each individual, how good are things overall – and all of the answers given here are logically compatible with relational egalitarianism.

<sup>10</sup>See Appendix 2 for two equivalent definitions of being spread-out: one in terms of mean-preserving spreads and one in terms of Pigou–Dalton transfers.

The final principle that prioritarrians accept is one that utilitarians accept and egalitarians reject:

NO DISTRIBUTIONAL GOOD: Relational or global properties do not have intrinsic value.

By ‘relational or global’ properties I mean properties that irreducibly concern more than one person. To adhere to NO DISTRIBUTIONAL GOOD is to hold that the only locations of intrinsic value are individuals, and that the only objects of intrinsic value are considerations pertaining to individuals. Properties of the distribution as a whole – like how much inequality there is – don’t have any independent value. NO DISTRIBUTIONAL GOOD rules out a particular explanation for SPREAD AVERSION. It rules out the egalitarian explanation that inequality is intrinsically bad (or equality is intrinsically good) – that it is in itself bad if some people are worse off than others, even if it isn’t bad for any particular person.

Call the conjunction of INDIVIDUAL GOOD, SPREAD AVERSION, and NO DISTRIBUTIONAL GOOD the *prioritarian commitments*.

Prioritarrians explain SPREAD AVERSION with reference to the fact that the worse-off should be given priority because they are the worse-off. But they also say something more specific: that those who are worse off *in an absolute sense* should be given priority (Parfit 1991: 23). We can see the primacy of absolute utility level by looking directly at the prioritarian aggregation rule  $P$ . The moral value function  $v$  transforms absolute utility level, independent of any distributional facts. How morally valuable it is to raise someone’s utility by a given amount depends on her absolute utility level, not on how much utility she has relative to someone else. Raising an individual’s utility level from, say, 100 to 200 has the same moral value whether she is the worst-off individual in a particular distribution or the best-off.

Thus, in addition to the three principles just mentioned, prioritarrians accept:

ABSOLUTE PRIORITY: It is more valuable to increase the utility of the absolutely worse-off than that of the absolutely better-off, regardless of their utility relative to others in the society.

For this reason, I will refer to prioritarianism as *absolute prioritarianism*. (Absolute prioritarianism should not be confused with the claim that the worst-off individual gets ‘absolute’ priority, in the sense that his interests trump everyone else’s. It means that priority depends on absolute utility.)

#### 4. Rank-weighted utilitarianism

We’ve just seen that absolute prioritarianism is an appealing way to satisfy the prioritarian commitments. It satisfies them by maintaining the ‘simple average’ or ‘simple sum’ structure of utilitarianism, but averaging or summing over moral value rather than prudential value. It thus commits itself to ABSOLUTE PRIORITY.

However, there is an alternative way to satisfy the prioritarian commitments. I will show this by first introducing a formal rule and then giving it an

interpretation that adheres to the three principles that make up the prioritarian concerns, as well as a fourth principle that I term RELATIVE PRIORITY. In the next section, I will argue that one can indeed hold all four principles on the basis of a single idea about value. (That the first three have been thought to rule out taking account of the relative position of the individuals in question is why, I think, this alternative has been overlooked philosophically, even though the formal rule hasn't been.<sup>11</sup>)

The formal principle in its 'total' version is known as *rank-weighted utilitarianism* or the *Gini family*: transform the *weight* that each utility value gets in the total, so that the well-being of relatively worse-off individuals gets more weight and the well-being of relatively better-off individuals gets less weight. Specifically: rank the individuals from worst-off to best-off, where individuals who are tied can be put in any order. For example, in D1, Bob is 1, Ann is 2, Cecil is 3; in D2, we can put them in any order. Each rank gets a weight  $\lambda$ , where the weights are larger for lower ranks. Multiply each utility by the rank-weight and sum the results. For example:

$$\lambda_1 = 5, \lambda_2 = 3, \lambda_3 = 1$$

$$W_{total}(D1) = 5(100) + 3(200) + 1(300) = 1400$$

$$W_{total}(D2) = 5(200) + 3(200) + 1(200) = 1800$$

In the 'average' version, assign weights to rank-ordered *proportions* (e.g. to the top 1/3 of individuals) rather than individuals, so that the weights sum to 1, and the weights of better-off groups of a given size are lower than the weights of worse-off groups of that size.

There are two equivalent ways to do the calculation for the 'average' version. The first: fix an *importance function*  $I(p)$  from  $[0, 1]$  to  $[0, 1]$ , measuring the importance of the interests of the best-off group of each size  $p$ . Let  $I$  be convex: the best-off group is less and less important as its size gets smaller and smaller. (For example, the group of everyone gets importance 1; the best-off 2/3 of individuals get importance 4/9; and the best-off 1/3 of individuals get importance 1/9.) Then let the *weight*  $w$  of

<sup>11</sup>A recent exception is Hirose (2015), who discusses and ultimately endorses the formal view I argue for here, which he calls 'the aggregate view of telic egalitarianism'. His discussion focuses on the differences between the form of  $E$  and the form of  $W$  (to be introduced in this section), where each is given the ethical interpretation its form naturally suggests. When he compares the formal view to absolute prioritarianism, he focuses on different implications of each view and thus compares them in terms of the palatability of their implications. Here, by contrast, I give positive philosophical reasons that one should accept the view (i.e. reasons one should care about relatives in the way I argue for here), independently of the other views on offer. That is, I give the formal view and a philosophical interpretation of it. (One can find in various places in Hirose, suitably interpreted, support for the idea that it is possible to accept  $W$  and all four principles articulated here, though, again, he does not give a positive reason for accepting this package of things.) A related view is Persson's (2001, 2008, 2011, 2012) *Relational* form of the Priority View, which he describes as a 'form of egalitarianism which is equivalent to a form of prioritarianism other than Parfit's form' (Persson 2008: 296) and which holds that 'instead of letting the absolute welfare level of recipients weight benefits (and harms) ... the relation of (un)just (in)equality weight[s] them' (Persson 2001: 35). The primary difference between Persson's relational view and relative prioritarianism is that the relational view accepts that (un)just inequality has intrinsic disvalue, thereby denying NO DISTRIBUTIONAL GOOD.

each rank-ordered group be importance of those at least as well off as that group minus the importance of those better off. This will mean that groups at the top get less weight than their size, and groups at the bottom get more weight than their size. Finally, multiply each utility by the weight of the group that obtains it. For example:

$$I(\text{top } 1/3) = 1/9, I(\text{top } 2/3) = 4/9, I(\text{top } 1) = 1$$

$$w(\text{top } 1/3) = 1/9, w(\text{middle } 1/3) = 4/9 - 1/9 = 3/9, w(\text{bottom } 1/3) \\ = 1 - 4/9 = 5/9$$

$$w(\text{top } 1) = 1$$

$$W_{\text{average}}(D1) = (5/9)(100) + (3/9)(200) + (1/9)(300) = 156$$

$$W_{\text{average}}(D2) = (1)(200) = 200$$

Thus, the utility values of relatively worse-off individuals get higher weight, and the utility values of relatively better-off individuals get lower weight.

The second, equivalent way to do the calculation is this. The top  $p$ -portion of individuals gets a particular weight – the importance value  $I(p)$  of this group, as above – and these weights attach to utility *increments* – utility that those in the top  $p$ -portion get but that those below don't. For example, utility that everyone receives gets weight 1; utility that the top 2/3 receive over and above what everyone receives gets weight 4/9; and utility that the top 1/3 receive over and above what the top 2/3 receive gets weight 1/9:

$$I(\text{top } 1/3) = 1/9, I(\text{top } 2/3) = 4/9, I(\text{top } 1) = 1$$

$$W_{\text{average}}(D1) = (1)(100) + (4/9)(200 - 100) + (1/9)(300 - 200) = 156$$

$$W_{\text{average}}(D2) = (1)(200) = 200$$

I invite the reader to keep in mind whichever equation is more intuitive. The first equation concerns how much weight to put on the fully described consequences belonging to each individual: e.g. 'those in the top 10% but not the top 5% have lives that look like this – how much does that fact contribute to the overall good?'. The second equation concerns how much weight to give to advantages that those in one group have but those in another do not: e.g. 'the top 10% of people enjoy at least 10 extra years of life in addition to those things enjoyed by those less fortunate – how much does that fact contribute to the overall good?'. The differences between them won't matter to the discussion here.

On this picture, D2 is better than D1, because D2 has a higher *rank-weighted average* or *rank-weighted total* utility, where those with relatively lower utility values are given more weight. When a distribution is more spread out in terms of utility and we put more weight on relatively worse-off individuals, individuals with lower utility bring down the rank-weighted average utility more than individuals with higher utility bring it up. Thus, if relatively worse-off individuals get more weight, then when two distributions have the same average

utility, the one that is less spread out in terms of utility will have a higher rank-weighted average.<sup>12</sup>

This aggregation rule gives priority to those who are worse-off in a *relative* sense: the weight given to those at the bottom of a particular distribution is higher than the weight given to those at the top of that distribution. The welfare of those who are relatively better off makes less of a difference to the overall good than the welfare of those who are relatively worse off.

$W$  (rank-weighted utilitarianism) is a sort of dual to  $P$  (absolute prioritarianism): instead of transforming the utility values (with  $\nu$ ) and keeping group sizes as weights, we transform the weights (with  $I$  or  $w$ ) and keep utility values. It is unfortunate that  $\nu$  has been sometimes called an (absolute) prioritarian ‘weighting’ function instead of a ‘value’ function, since it transforms utility *values*, not the weight that each utility value gets – and the key formal difference between absolute and relative prioritarianism is which function gets transformed. To maintain precision, I will refer to the absolute prioritarian  $\nu$  as a value function and the relative prioritarian  $I$  or  $w$  as a weighting function – more on the philosophical distinction between value and weight in subsequent sections.

$W$  and  $P$  do not give rise to the same orderings of social distributions: there are orderings that can be captured by  $W$  but not  $P$ , and vice versa.<sup>13</sup> So they are not merely notational variants of each other with different philosophical interpretations; instead, they say different things about which distributions are better than which others, and in what way the interests of the worse-off should matter more.

When weighted-rank utilitarianism is discussed in the philosophical literature, it is interpreted as an egalitarian principle, for reasons I will discuss shortly. But it needn't be. I will argue that a particular ethical interpretation of  $W$  is a plausible way to satisfy the prioritarian concerns: INDIVIDUAL GOOD, SPREAD AVERSION, and (perhaps surprisingly!) NO DISTRIBUTIONAL GOOD. This interpretation accepts these claims and, instead of ABSOLUTE PRIORITY, accepts:

RELATIVE PRIORITY: It is more valuable to increase the utility of the relatively worse-off than that of the relatively better-off, regardless of their absolute utility.

It is more valuable to increase the utility of the relatively worse-off, the view says, because the utility of the relatively worse-off *counts more* in determining the total good than the utility of the relatively better-off – it is *weighed more heavily*. The overall good in a society comes from one source – how things go for each individual – but some instantiations of this (some individuals) contribute to or reflect the overall good more than others do. ***How good things are for worse-off individuals in a distribution makes more of a per capita contribution to how good things are for the group, because their good more strongly determines***

<sup>12</sup>See Appendix 2.

<sup>13</sup>To see this, note that  $P$  adheres to strong separability and  $W$  violates it; see section 6 of this article.

**how good that distribution is than the good of better-off individuals does.**<sup>14</sup> This is the crux of the view.

Call this view the view I've just described *relative prioritarianism*. Relative prioritarianism is characterized by the formal principle *W* in conjunction with the philosophical view that how good things are for worse-off individuals in a group more strongly determines how good things are for the group than how things are for better-off individuals does.

It is easy to see that relative prioritarianism accords with SPREAD AVERSION – the claim that of two distributions with the same average utility, the one that is less spread out in terms of utility is better – and RELATIVE PRIORITY – the claim that it is more valuable to give a bit of utility to a group of a particular size than to give that same bit of utility to a better-off group of the same size.

A bigger puzzle is how it can make sense to take a relative position into account while still holding that relative position does not affect individual utility, nor are relational properties a separate source of value.

## 5. Assessment and weight in aggregation problems

Before showing that relative prioritarianism satisfies the remaining two claims, it will be helpful to make a distinction that arises in aggregation problems – problems of evaluating the whole from its constituent parts – more generally. The distinction is between which *attributes* of the whole matter to the evaluation, the *assessment* of each instance of an attribute, and the *weight* that each instance gets in determining the assessment of the whole.

Consider the problem faced by an engineer who wants to determine the strength of a chain composed of many links. Let us assume that we have two hypotheses for how the strength of a chain relates to the strength of its links: its strength might either be an average of the strength of its links, or it might be the strength of its weakest link. Both aggregate the same *attribute* (strength), and, furthermore, they agree about how to *assess* this attribute (how to measure individual chain strengths). But they differ in that they assign different *weights* to bearers of the attribute.

Or consider the problem faced by an Olympic committee that wants to determine how to calculate a gymnast's overall score, based on her scores for four events. Let's say that the committee members score a gymnast's performance on each event according to the gymnastic ability she displayed, and they agree that the overall gymnastic ability displayed by an athlete should be the average of her four event scores, but they disagree about how to measure gymnastic ability: some members of the committee think it should be determined by execution, while others think it should be determined by execution and difficulty. Here, the members agree about the *attribute* to be aggregated (gymnastic ability) and about the *weight* of each bearer of the attribute (i.e. each event); but they disagree about how to *assess* the attribute.

<sup>14</sup>Rawls's (1971) 'difference principle' (interpreted as 'maximin') could be thought of as an extreme version of this view: the overall good in a society is *only* a matter how good things are for the worst-off group. (Formally, set  $I(1) = 1$  and  $I(p) = 0$  for  $p \neq 1$ .) This isn't, however, the motivation he gives for it.

Finally, consider two deans evaluating philosophy departments. They agree about how to measure individual professors' research output (we should count publications), and about how to weight them in determining a department's overall research output (we should sum individual research outputs), but they disagree about how overall department strength relates to research output: one dean thinks department strength just is research output, whereas the other thinks it is research output combined with faculty diversity. Thus, they disagree about which attributes matter to overall department strength. Here, they agree about both *assessment* and *weight* of research output in determining overall department research output; but they disagree about whether this *attribute* is the only thing that matters in their evaluation of department strength.

Two important things to note. First, the *assessment* of an attribute remains the same whether it is part of an aggregation problem or not, and it is stable across different aggregation problems: the strength of an individual chain link doesn't change when paired with other links; the execution score of a routine doesn't change when paired with other routines. By contrast, the *weight* of an attribute is essentially a feature of aggregation itself: there isn't an answer to 'how much does this attribute contribute' apart from knowing 'contribute to what whole?'

Second, which *attributes* are to be aggregated naturally corresponds to what *intrinsically matters* in the evaluation. The dean that thinks diversity matters to a department thinks that it matters intrinsically: for him, it's not that diversity contributes to what 'really' matters, a department's research output; rather, it is a separate part of a department's strength. By contrast, the members of the Olympic committee and the engineer's two hypotheses agree about what intrinsically matters to their overall evaluation – gymnastic ability on individual events and individual link strength, respectively.

Utilitarianism, egalitarianism and absolute prioritarianism are all interested in assessing the overall goodness of a distribution. Utilitarianism and absolute prioritarianism agree that only one attribute intrinsically matters to this evaluation – the moral value of each individual's well-being – but they disagree about how to assess it: for the utilitarian, how morally valuable it is that a particular individual has a particular level of well-being is identical to how good that level is for her (its utility), and for the absolute prioritarian, these are positively related but not identical.<sup>15</sup> Thus, utilitarians and absolute prioritarians agree about the attribute question and disagree about the assessment question. (Their disagreement with each other is like that within the Olympic committee.) Egalitarians, by contrast, hold that two attributes intrinsically matter to the moral value of a distribution: each individual's well-being (or the moral value thereof, which is identical to utility) and (in)equality. Thus, utilitarians and egalitarians agree about the assessment question, and disagree about the attribute question. (Their disagreement with each other is like that between the two deans.)

---

<sup>15</sup>We could instead say they care about different attributes – well-being and moral value – but this seems to obscure their disagreement. Still, nothing in this paper hangs on which of these things we say. We could similarly interpret the egalitarian either as caring about inequality plus individual well-being or inequality plus the moral value of individual well-being.

Interestingly enough, though, all three views *agree* about the weighting question. **They all equate the overall value of an attribute with a simple average or simple sum of that attribute.** Prioritarians say that the moral value of an individual's utility is not identical to utility, and they average or sum moral value to get overall goodness; egalitarians say that individual well-being isn't the only thing that matters to overall goodness, but overall well-being is the average or sum of individual well-being.

By contrast, relative prioritarianism disagrees with all three positions about the weighting question. Like utilitarianism and absolute prioritarianism, it holds that overall goodness is determined by only one attribute, the moral value of each individual's well-being; and like utilitarianism and egalitarianism, it holds that the correct measure of this is utility, not some function of utility. But unlike all three theories, relative prioritarianism holds that relatively worse-off individuals get more weight in the evaluation of overall goodness than relatively better-off individuals do: how things go for the former is more important than how things go for the latter. Thus, utilitarians and relative prioritarians agree about the attribute *and* assessment question, and disagree about the weighting question. (Their disagreement with each other is like that between the two theories of chain strength.)

It is not that the relative prioritarian starts with average utility and then gives *additional* weight to those who are relatively worse off. To say this would be to imply that 'overall' *means* simple average or simple sum. On the contrary, *the claim that 'overall' should be analyzed as 'simple average' or 'simple sum' is a substantive view, not a logical equivalence or even a default assumption.* (It is not analytic, for example, that the 'overall' strength of a chain is the average or sum of its link strengths.) Utilitarianism and relative prioritarianism both rank distributions according to their overall utility – but they make different substantive commitments about what 'overall' means.

## 6. Vindicating INDIVIDUAL GOOD and NO DISTRIBUTIONAL GOOD

Relative prioritarianism gives more weight to those who are relatively worse off than to those who are relatively better off. In order to see what this implies about individual and distributional good, it will be helpful to look at some other examples in which instantiations of an attribute are given more weight because of their relative position. It is the basis for a cliché that a chain is only as strong as its weakest link.<sup>16</sup> It is sometimes said that a mother is only as happy as her least happy child. In many Olympic events (e.g. downhill skiing, long jump, pole vault, shot put), one's time or distance for an event is equal to one's best time or distance out of a number of trials. In all of these cases, the members of a set of things (links, children, trials) each possess an attribute (strength, happiness level, time or distance), but certain members are given more weight: the members contribute differently to the attribute of the whole, depending on their relative position (weakest, least happy, fastest time or farthest distance).

<sup>16</sup>This may be technically false, but we can understand the claim it is making about chains.

These are examples in which only the extremum gets any weight at all, but that is simply a limit case of weight based on relative position.<sup>17</sup> We could instead define a jumper's overall score to be two-thirds the distance of his best jump plus one-third the distance of his second-best jump: multiple members of the set of jumps contribute something to the overall score, and the weight of each jump depends on how long the other jumps are.

With these facts in mind, I will now show that relative prioritarianism is compatible with INDIVIDUAL GOOD and NO DISTRIBUTIONAL GOOD.

Let me start by showing that relative prioritarianism is compatible with holding that individual good is independent of relative position (i.e. with INDIVIDUAL GOOD). Consider the examples of chain strength, parental happiness and Olympic event score. In these examples, a relative fact determines how a particular link's strength, a particular child's happiness, or a particular jump's distance affects overall strength, happiness and score: how strong, happy or far that entity is relative to other children, jumps or links. But that does not mean that one link's strength is partially constituted by another link's strength, one child's happiness is partially constituted by her sibling's happiness, or one jump distance is partially constituted by another's distance. In other words: one can give more *weight* to an instantiation of an attribute, based on its relative position, without changing the *assessment* of that attribute.

Similarly, relative prioritarianism holds that the good of the worse-off counts more heavily in the overall good than that of the better-off counts, without holding that the good of each individual is partially constituted by whether she is worse-off or better-off than her fellows. A particular level of well-being has a fixed value no matter whether the person who obtains it is the worst-off person or the best-off person: one's *utility* doesn't go down if someone else enjoys a higher utility level. Rather, how much that person's utility *counts* towards the total depends on which relative position he occupies: his utility matters more to the overall value of the distribution when someone else enjoys a higher utility level.

Next I will show that the relative prioritarian thought – that the good of the relatively worse-off counts more – is compatible with holding that distributional or relational facts do not provide an intrinsic source of value (i.e. is compatible with NO DISTRIBUTIONAL GOOD).

There are two ways in which an aggregation rule might be sensitive to spread. First, spread can be a separate consideration, to be thought of alongside, and weighed against, facts about the good of individuals. Second, spread can fail to be a separate consideration, but nevertheless it may follow from the way utility is aggregated that of two distributions with the same overall utility, the one that is less spread out in terms of utility is better – it may follow merely as an epiphenomenon of the aggregation rule. Relative prioritarianism takes the latter to be true. (So does absolute prioritarianism: both are ways of saying that less-spread-out utility distributions are better, but not because utility spread is bad in itself.) To return to our analogies: if the cliché is true, then holding fixed the average strength of a chain's links, the chain whose strengths are less spread-out

<sup>17</sup>Thus, these examples all map on to maximin, which is a limit case of relative prioritarianism (see footnote 14).

will be stronger, but not because *spread* of link strengths is a consideration in overall chain strength; instead, because for a given average, less-spread-out means that the minimum is higher. And similarly for spread in happiness of children or event scores – a mother doesn't care about how spread out her children's happiness levels are, and an Olympic committee doesn't care about how spread out a jumper's distances are, but averages being equal, less-spread-out happiness levels will be better and more-spread-out distances will be better.<sup>18</sup>

The fact that the utility of the worse-off contributes more to the overall value of the distribution than the utility of the better-off also does not mean that the relative standing of an individual is a separate attribute to be valued, like (in)equality is for the egalitarian. Again, even for a fixed view of which attributes matter and how to assess them, there is a substantive question of how to weight each location of the attribute in determining the whole. In telling us how much weight to give to each person's utility, relational considerations tell us how locations of intrinsic value should be combined – they do not *have* intrinsic value.

Thus, relational considerations can partially explain the value of a distribution without either affecting an individual's well-being or constituting a separate source of value. Instead, they determine *how much* each individual's well-being contributes to overall goodness.

To summarize. Relative prioritarianism holds that in terms of individual utility, it does not matter how one ranks relative to others (INDIVIDUAL GOOD holds). How individuals rank relative to each other does matter to the overall goodness of a distribution, but not because distributional or relational facts are themselves valuable; instead, relational facts determine how much each individual's utility contributes to overall good (NO DISTRIBUTIONAL GOOD holds<sup>19</sup>). And since relatively worse-off individuals count more heavily in the overall good, less-spread-out distributions are better, when their simple averages are equal (SPREAD AVERSION holds).

The reader might wonder: is the distinction between valuing distributional facts in themselves and holding that they determine how the actual sources of value combine an important distinction? There are two objections here. First, one might hold that *rankings* of social distributions are all there is, and so two views which produce the same rankings are 'the same' view. Second, and relatedly, one might object to the particular way I've been characterizing egalitarianism (as the

<sup>18</sup>These points are easier to see if we care more but not exclusively about the extremum of a set; for example, if parental happiness doesn't depend entirely on the least happy child, but instead depends more on unhappier children than happier children.

<sup>19</sup>An argument that reaches a similar conclusion is in Hirose (2009). Hirose is interested in egalitarianism, interpreted as any formalism which ranks distributions equivalently to some version of *E*. He notes that *W* fits this criterion. He then points out that we can adhere to *W* while thinking that "Inequality is not an object of aggregation for estimating the goodness of a state of affairs, but a feature of an aggregative process for estimating the goodness of a state of affairs" (Hirose 2009: 303). He also notes that, on *W*, "the relation between different people determines how much priority (or moral importance) we give to each person's wellbeing" (308). In our terms, therefore, he argues that one could accept *W* by accepting RELATIVE PRIORITY rather than by denying NO DISTRIBUTIONAL GOOD. He does not, however, explain what might motivate this choice (that is not the point of his argument). This argument is reconstructed in Hirose (2015: 74–78), where he also notes some objections to thinking of inequality as bad in itself.

rejection of NO DISTRIBUTIONAL GOOD), and prefer to instead define it by a particular formal principle that I will mention shortly.

One motivation for both of these objections is the realization that both the absolute prioritarian formalism  $P$  and relative prioritarianism formalism  $W$  can be rewritten as special cases of the egalitarian formalism  $E$ : they can be separated into a simple utility average and a remainder ('penalty for inequality').<sup>20</sup> (Simply define the remainder as the difference between the output of a utilitarian aggregation function and the output of the aggregation function in question.) More generally: any two aggregation views can be made to have the same assessment and weighing of a particular attribute, by holding that one of the views cares about an additional attribution as well, as long as that attribute is allowed to be defined in a non-natural way.

Since a ranking of social distributions can result from multiple different formalisms, it appears that rankings are the only 'real' thing – the first objection. Furthermore – the second objection – given that  $P$  can be seen as a special case of  $E$ , some philosophers have been interested in distinguishing prioritarian rankings (rather than prioritarian aggregation rules) from egalitarian rankings. These philosophers have settled on holding that prioritarian rankings are those that are spread averse and *obey* a principle called strong separability, whereas egalitarian rankings are those that are spread averse and *violate* strong separability.<sup>21</sup> Strong separability says that to figure out whether it is better (increases the good more) to increase Ann's utility by some amount or increase Bob's utility by some amount, we do not need to look at what Cecil has – or, put another way, whether it is better for Ann and Bob to have some utility amounts or for them to have some different utility amounts has the same answer regardless of what Cecil has.<sup>22</sup>

Let us begin with the first objection: that only rankings, not aggregation rules, are real. Since the relative prioritarian ranking can be derived from a rule that is a combination of a simple average and a penalty for inequality, the objection goes, the distinction between the view that inequality is in itself bad (egalitarianism) and the view that the good in a society is more sensitive to the good of the relatively worse-off (relative prioritarianism) is a distinction without difference.

In reply, it's true that these views have the same upshots about which social distributions are better than which other social distributions. However, they have different *motivations* for these upshots, which is important. Arguments in favour

<sup>20</sup>For  $P$ , see Jensen (2003) and Fleurbaey (2015). If one conceives of absolute prioritarianism in this way, it is vulnerable to a kind of levelling-down objection, since levelling-down can decrease the remainder (see Persson (2008, 2011, 2012) for this objection, and Holtug (2010: 210), Porter (2010), and Weber (2019) for replies). For  $W$ , see Weymark (1981) and Hirose (2009). In the case of  $W$  with coefficients that are the first  $n$  odd numbers renormalized, the 'penalty for inequality' is given by the Gini coefficient, a well-known measure of inequality (see Weymark 1981: 412). In my opinion, this is largely why  $W$  has been thought of as an egalitarian principle in the philosophical sense. (See Blackorby and Donaldson (1978) for a more general result about the relationship between aggregation rules and indices of inequality.)

<sup>21</sup>See e.g. Broome (1991), McCarthy (2008). Jensen (2003) holds that egalitarianism should be identified with the denial of additive separability. Additive separability implies strong separability, and the converse holds under certain conditions (see Broome 1991: 82–86).

<sup>22</sup>See e.g. Broome (1991: 69) for a formal statement of strong separability.

of one or another view in distributive ethics do not simply consist in lists of rankings that are self-evidently better or worse than other lists of rankings. They also (and in my opinion more centrally) involve giving a compelling view about what sorts of things are valuable. To say that equality is valuable is to say something different – something that requires different arguments and gives rise to different objections – than to say that overall good is determined more by how good things are for the worse-off than for the better-off.<sup>23</sup> Those who have been moved to dismiss egalitarianism – and thus the ranking it gives rise to – because they hold that inequality is not intrinsically bad may be willing to accept relative prioritarianism – and thus accept the (identical) ranking it gives rise to.

Let us turn to the second objection: that egalitarianism is (or should be defined as) simply SPREAD AVERSION plus a denial of strong separability, and so I haven't introduced a view distinct from egalitarianism. Call the definition of egalitarianism in terms of holding that there is distributional good *philosophical egalitarianism*, and the definition in terms of denying strong separability *formal egalitarianism*. My main argument will still be of interest to those who prefer the formal definition. For it shows that it is possible to be a formal egalitarian without being a philosophical egalitarian. In particular, it shows us that there are two different reasons one might reject strong separability, i.e. two different reasons one might hold that what Cecil has matters to whether it is better for Ann and Bob to have some utility amounts or for them to have some different utility amounts. The first reason is the 'philosophical egalitarian' reason: what Cecil has partially determines the contribution of Ann's and Bob's utility to overall inequality, and inequality is bad. But the second, newly noticed, reason is the relative prioritarian reason: what Cecil has partially determines where Ann and Bob are in the relative ordering, and thus how much weight Ann's and Bob's good each get in determining the total good. For example, if Ann is the worst-off and Bob is the best-off, then Ann's good matters a lot more than Bob's, but if Ann and Bob are the two worst-off, then while Ann's good still matters more than Bob's, it does not matter quite as much more.

We have seen that relative prioritarianism is a coherent option available to those motivated by the prioritarian commitments. In the next section, we will see what there is to recommend relative prioritarianism over absolute prioritarianism.

## 7. Relative priority vs. absolute priority

If one is motivated by the prioritarian commitments, what is there to distinguish between absolute and relative prioritarianism? The main reason to choose one of these options is what we think constitutes overall good.

It will be helpful to keep in mind four distinct concepts, which follow from distinctions made above. The *prudential value* to an individual of being at a certain utility level is just that utility level. The *moral value* of an individual being at a certain utility level is the value of that individual being at that utility

---

<sup>23</sup>For example, while the 'levelling-down' objection is a potential objection to the view that equality is good in itself, this objection is an obvious non-starter against the view that overall good is determined more by how things are for the worse-off. See also Hirose (2009) for a similar point.

level, from a moral point of view, for example, from the point of view of a morally motivated individual who needs to decide how to divide goods among strangers. The *prudential weight* of the prudential value associated with an individual's consequence is how much weight that value gets in determining which distribution is prudentially better for her; under certainty, the prudential weight of the individual's actual consequence is always 1 (under uncertainty, for an expected utility maximizer, the prudential weight of a consequence is equal to the probability of that consequence). The *moral weight* of the moral or prudential value associated with an individual's consequence is how much weight that value gets in determining which distribution is morally better. Using terms from section 5: prudential value is the attribute we aggregate to determine what's good for an individual, and prudential weight is the weight we give to each instantiation of prudential value; moral value is the attribute we aggregate to determine what's morally good, and moral weight is the weight we give to each instantiation of moral value.

Aggregative axiology asks: if two distributions differ only in their utility for a single individual (say, Ann gets 110 utils rather than 100), how morally better is the one than the other, and how is this determined? In answering this question, there are two things to be assessed.

First, how morally valuable is it that Ann has 110 rather than 100? Relative prioritarianism, like utilitarianism, says that it is as morally valuable as it is prudentially valuable for Ann (10 units of value). By contrast, absolute prioritarianism says that how morally valuable it is depends on how much prudential good Ann already has (the more prudential good Ann already has, the less morally valuable it is).

Second, how much moral weight do we give to this fact – the fact about how morally valuable it is that Ann has one consequence rather than another – compared with the facts about how morally valuable it is that other individuals have what they have? Absolute prioritarianism, like utilitarianism, says that the moral value of Ann having what she has gets identical moral *weight* to the moral value of others having what they have. By contrast, relative prioritarianism says that the moral value of Ann having what she has gets less moral *weight*, the better off she is compared with everyone else. (To stress the point made earlier: this is different from saying that differences in Ann's utility level have less moral *value* – getting less weight and having less value are both ways for changes in Ann's utility to matter less to the *overall* value of a distribution, but they are different.<sup>24</sup>)

If we are motivated by the prioritarian commitments *and* think that incremental increases in Ann's prudential good become less morally valuable the more she has of it, then we should be absolute prioritarians; but if we are motivated by the

<sup>24</sup>There are two ways that a jumper's wearing lighter shoes on some particular trial might not matter to his final score: it might be that wearing lighter shoes does not increase his distance on that particular trial (it doesn't change the relevant value), or it might be that the particular jump, by virtue of being the shortest of his jumps, does not count towards his overall score (whether the relevant value changes, the weight is zero). These are both ways for a jumper's change in footwear to not matter to his overall score, but they are different.

prioritarian commitments *and* think that the value of what Ann has gets less moral weight in the overall distribution the more she has relative to others – that Ann’s well-being is less important to the overall good than the well-being of worse-off individuals – then we should be relative prioritarrians.

Why does the relative prioritarian think that the well-being of worse-off individuals is more important (counts more) to the overall good than the well-being of better-off individuals? She thinks that the good of the worse-off *better reflects* the overall good than the good of the better-off.

For certain Olympic events, we hold that an athlete’s prowess is reflected better by his *best* distance or time than by his *average* distance or time; we could instead have decided that his prowess is reflected by an average of his times, but presumably we think that there is some conceptual link between athletic prowess and the best one can achieve, at least in these events. (In the modified example, in which an athlete’s score is two-thirds his best distance plus one-third his second-best distance, we would think there is some conceptual link between athletic prowess and the best one can achieve, and also some conceptual link between athletic prowess and the second-best one can achieve – but that the former link is stronger.) For a chain, there is (assumed to be) a physical link between the strength of the weakest link and the overall integrity of the chain. For a parent’s happiness, we might think there is a causal link between the happiness of the least happy child and the parent’s happiness.

The relative prioritarian holds that there is a conceptual link between the goodness of a society and how its worse-off members fare; there is also a conceptual link between societal goodness and how its better-off members fare, but the conceptual link to the worse-off is stronger. (There is not, or at least not directly, a conceptual link between societal goodness and the average of how its members fare.)

So, the main reason to choose relative prioritarianism over absolute prioritarianism is that one thinks that the former correctly tracks the facts about the overall good and captures the sense in which we should give priority to the worse off. The relative prioritarian makes two separate claims: that moral value should be prudential value, and that the worse-off should get more moral weight than the better-off. And she justifies the latter claim by holding that there is a strong conceptual link between the goodness of a society and the well-being of its worse-off members, stronger than the conceptual link between the goodness of a society and the well-being of its best-off members.

### 7.1. Problems with absolute prioritarianism

The absolute prioritarian’s claim that moral and prudential *value* come apart gives rise to three concerns, that do not arise from the relative prioritarian’s claim that moral and prudential *weight* come apart. (None of these worries is unique to me, but there is some point to gathering them together in the same place and showing that relative prioritarianism doesn’t face them.) The first concern arises in evaluating a single-person case. If moral and prudential value come apart, then it looks like a morally motivated stranger should choose differently on a subject’s behalf than

the subject herself would choose, even if the morally motivated stranger is motivated only by concerns involving her well-being.

To make this concrete, assume Ann is the only person in the world, and she has (by definition) equally strong prudential reasons for going from utility 100 to utility 200 as she does for going from utility 200 to 300. Surely a morally motivated stranger – motivated only by what’s good for Ann – should have equally strong moral reasons to help her go from 100 to 200 as he does to help her go from 200 to 300; but absolute prioritarianism says that he has stronger moral reasons in the former case.

This is a worry about *conceptual* machinery of absolute prioritarianism: we must accept something counterintuitive about the strength of our reasons for helping others, versus the strength of their reasons for them helping themselves. But it also gives rise to a potential *inconsistency*, if we hold that when evaluating distributions under risk, we should first determine the prioritarian value of each distribution and then take risk into account (the *ex post* view).<sup>25</sup>

According to a common view of rationality, a person’s prudential value under risk is given by the expectation of his utility function; so if, for example, Ann is choosing between 200 utils and the fair-coin flip {HEADS, 100; TAILS, 300}, she should be indifferent; and if a very small amount of utility is added to the coin-flip, she should choose it. But now consider the case in which I am making the choice for Ann, and I am concerned with maximizing moral value. *Ex post* absolute prioritarians apply the moral value function to utilities of consequences (rather than risky prospects) and thus hold that the moral value of a distribution is its expected moral value. As long as the value function is concave enough, I should choose the sure-thing, even with the ‘sweetener’ added to the coin-flip.<sup>26</sup> So I should choose against Ann’s self-interest, even when she is the only person affected. (The same point holds if one opts for a different view of rational prudence than expected utility maximization, as long the view doesn’t itself adopt a prioritarian transformation of utilities.)

This problem does not plague *ex ante* absolute prioritarianism, which applies the moral value function to *expected* utilities; nor does it plague *factualist* prioritarianism, which applies the moral value function to the utilities that *in fact* result from a prospect.<sup>27</sup> The *ex ante* view will order Ann’s options as she does, and the factualist view will order Ann’s consequences given what in fact happens as she does. However, they will not assign the same numbers to Ann’s options or consequences as she does: these numbers will be a concave transformation of hers. Thus, while these views won’t posit inconsistent evaluations, they still face the conceptual problem of explaining how prudential and moral value are meaningfully different in the single-person case.

<sup>25</sup>Otsuka (2015). A similar type of argument appears in Broome (1991: 217). (See Otsuka and Voorhoeve (2009) for related worries.) Otsuka points out that this worry also applies to a ‘hybrid’ view suggested by Parfit.

<sup>26</sup>For example, using our value function above,  $v(200) = 190$ , and  $v(\text{coin-flip}) = 0.5(100) + 0.5(250) = 175$ , so as long as the ‘sweetener’ is small enough, the sure-thing has higher expected moral value.

<sup>27</sup>For a defence of the latter, see Holtug (2019).

A second worry stems from a technical problem. Versions of this point have been raised by both Broome (1991) and McCarthy (2013, 2015). The basic worry is that in a risk-free framework, the absolute prioritarian can't distinguish, on the basis of an ordering of distributions, whether a particular distribution is better because of the shape of the utility function or the shape of the moral value function. For example, if we think that the distribution in which both Ann and Bob have a moderately happy life is better than the one in which Ann has a very happy life and Bob has an unhappy life, is this because the utility difference between having a very happy life and having a moderately happy life is smaller than the utility difference between having a moderately happy life and having an unhappy life, or is this because the *moral value* difference in *utility* between the former is smaller than that between the latter? Mathematically, the problem is that when goodness is determined by the average or total of  $v$  composed with  $u$ , we have no way to separately determine  $v$  and  $u$  from the overall values of distributions or from their ranking – many pairs of  $v$  and  $u$  will yield the same averages. Thus, without having an independent way to determine utility (independent of which distributions we think are better), there is no way to separate utility and moral value.

I mentioned earlier that we are not just concerned with orderings, but also with the reasons behind them. So simply saying that the utilitarian could reproduce any prioritarian ordering (and vice versa) does not constitute an objection. However, the point goes deeper than this. If we have two explanations for a given ordering, one which holds that moral good and prudential good are the same, and the other which holds that they are different, then we had better have some independent access to prudential good in order to argue for the latter. Without this independent access, one cannot argue against utilitarianism on the grounds that utilitarianism fails to account for inequality-aversion. (Another way to put the point: if one is looking to explain the intuition that utilitarianism fails to account for inequality-aversion, then one must either claim we have independent knowledge of prudential good, or explain the intuition using a theory other than absolute prioritarianism.)

The third worry is related. We need some idea of how prudentially good a consequence is in an absolute sense not only to distinguish between utilitarianism and absolute prioritarianism, but to *apply* absolute prioritarianism. To apply absolute prioritarianism we need there to be a meaningful zero-utility point.<sup>28</sup> But, first, it is unclear that such a thing has meaning. Second, even if it does and there is a meaningful point, it seems difficult to determine. Third, even if we could determine it, we *already* seem to have access to the relevant moral facts that help us intuit which distributions are better than which other ones without doing so – so facts about the zero point do not appear to be the relevant facts in our ordinary moral judgement. Finally, this cuts off one possible response to the second worry, which is to hold that our independent knowledge

---

<sup>28</sup>Different absolute prioritarian rules require different meaningful facts. For example, the Atkinson social welfare function is invariant under ratio-rescaling, meaning that the zero point is meaningful but multiplying the utility function by a positive constant will not change the ordering of distributions. Other absolute prioritarian rules require more information; see Adler (2012), especially the chart on p. 312.

of prudential good arises via a von Neumann–Morgenstern utility function, since a vNM utility function does not give rise to a meaningful zero-point.

It might be that these worries can be overcome; but ways of overcoming them bring their own costs.<sup>29</sup> And, importantly, relative prioritarianism does not suffer from any of these worries. It is conceptually easy to distinguish between the relative size of a group and that group's relative weight in assessing the overall good. Moral value always coincides with prudential value; and in the single-person case, the person's prudential weight of her own interests will be the same as the moral weight of those interests (i.e. 1). So a morally motivated stranger has just as strong reasons to help an individual person as this person has to help herself, in the single-person case. Furthermore, both *ex ante* and *ex post* views will produce an ordering that coincides with the individual's prudential ordering, and moral values that coincide with prudential values. We can distinguish, from a given ordering of distributions, between the group size, the weight of each group, and the utility of each consequence.<sup>30</sup> Finally, we only need to know facts about the relative value of consequences.<sup>31</sup> And it seems that we do know these facts; indeed, these are the facts that appear to drive our intuitions about what to do. As Larry Temkin points out, we are moved by the plight of a typical poor person in the United States because of how she fares relative to others in her society, not because of how well-off she is in an absolute sense, because by historical standards she is actually very well-off (Temkin: 2003b: 70–71).

The absolute prioritarian's claim that moral and prudential value come apart may raise additional worries for anyone motivated not just by the idea that only personal facts have intrinsic value (NO DISTRIBUTIONAL GOOD) but also that all value must be *personal*: value must be value *for someone*. Egalitarianism holds that there is impersonal value, namely, the (dis)value of (in)equality. Absolute prioritarianism also appears to hold that there is impersonal value, since moral value isn't value for anyone (this is what it is for  $v$  and  $u$  to come apart): there is *something* good about someone's increase in welfare that is not good for her

<sup>29</sup>Williams (2012) seeks to overcome the first worry by arguing for a restricted version of absolute prioritarianism, according to which in the individual case we should maximize expected utility but in the group case we should apply the absolute prioritarian value function. (Williams argues for a deontic version of this view; an axiological version would have to say that the impersonal good of a person's having a good depends on their absolute good, but only when others are involved.) The cost of this view is less parsimony. Adler (2012: 343–344) seeks to overcome the second and third worries by arguing that we can derive individual utility from extended preferences among lotteries over life-histories, with utility 0 assigned to non-existence (i.e. he introduces a framework with risk and a commitment to a zero-point). The costs of this view are whatever commitments are involved in assigning utility in this way, plus accepting that we need a framework with risk in order to understand distributive ethics.

<sup>30</sup>To see this, simply adapt the representation theorem in Buchak (2013) to be about social choice rather than decision theory, i.e. interpret  $p$  as proportions rather than probabilities,  $E$  as groups of people rather than events, and  $r$  as the importance function  $I$  rather than a risk function. Notice that even if  $p$  is not given, we could derive  $p$  from the ranking of social distributions. This is because  $p$  and  $I$  correspond to different features of the rankings (see Buchak (2013: Ch. 3) for an explanation of this in the decision theory case).

<sup>31</sup>See also Hirose (2015: 100). Note that if we want to read the relative values of consequences off of vNM utility functions, this does require interpersonal comparability, but not a meaningful zero-utility point.

or anyone else.<sup>32</sup> By contrast, relative prioritarianism holds that all value is personal – the only thing that is valuable is a person’s well-being. (One might object: don’t the weights that attach to each person’s well-being imply a commitment to impersonal value, namely the contribution of each person’s well-being to the overall good? No. Weights are merely the extent to which each personal value figures into the overall good. All of the views mentioned here, even utilitarianism, is committed to weights – they are typically just equal weights – so relative prioritarianism is committed to impersonal value in virtue of giving more weight to the worse-off only if the other views are also committed to impersonal value in virtue of giving equal weight to the worse-off.)

### 7.2. Responses to worries about relative prioritarianism

Might relative prioritarianism suffer from different worries, from which absolute prioritarianism is immune? One worry is that giving more weight to the worse-off amounts to thinking that some people matter more than others; another is that failing to take account of absolute position is morally objectionable. I will consider each in turn.

The first worry is that by weighting some people’s interests more strongly than others, we thereby hold that some people matter more than others, contrarily to a commitment to treating people the same. One might think: to treat people the same is to weigh their interests the same, that is, to give facts about each person’s well-being identical weight in determining the overall good. In reply, there is a weaker sense of treating people the same which I think better captures the morally relevant concept: to take someone’s well-being into account in the same way whether that someone is Ann, Bob or Cecil. For example, the distribution {Ann, 200; Bob, 100; Cecil, 300} should be just as good as the distribution {Ann, 100; Bob, 300; Cecil, 200}. This sense of same-treatment is often referred to as *anonymity* – the identities of the individuals can be swapped without changing the overall value – and all of the views in this paper (including relative prioritarianism) adhere to it. I contend that anonymity, not the stronger identical-weight principle, actually captures the kind of same-treatment ethicists should be concerned with. If this is right, then how much weight to give to each individual cannot be settled by considerations about treating people the same: all of the views here treat people the same in this sense.

Let us move on to the second worry. Here is a standard sort of intuition that seems to support absolute prioritarianism. Improvements to the utility of someone who is at a very low absolute utility level because she is constantly hungry or in excruciating pain should weigh more heavily than improvements to the utility of someone who is at a much higher absolute level, both in the single-person case – we have more reason to help the person in the first scenario than in the second – and in a multi-person case – we have even stronger reason to

---

<sup>32</sup>See discussion in Persson (2001, 2008, 2011, 2012) and Holtug (2010), particularly about the levelling-down objection to absolute prioritarianism.

help the person over others in the first scenario than we have to help the person over others in the second scenario, even if these others are all better-off.<sup>33</sup>

These intuitions are compelling. However, we have to be careful to distinguish between the idea that some given change I can make in the world (say, giving someone a sandwich) provides more utility when the recipient is at a low utility level, and the idea that some given *increment of utility* provides more moral value. The relative prioritarian will, of course, say that it is more morally valuable to give a sandwich to someone when he suffers from constant hunger than it will be to give him a sandwich when he has enough to eat: the former increases his utility much more (it is more prudentially valuable).

To make trouble for the relative prioritarian, one must say that it is more *morally* valuable than it is *prudentially* valuable to improve the lot of the person at the low level. (Or that it is less morally valuable than it is prudentially valuable to improve the lot of the person at the high level.) Focus on the difference that having a daily sandwich makes to a person with constant hunger – it does a lot to improve his well-being. Now imagine what change would have to take place in the world to improve this person's well-being by the same amount, if he started out in the position of having enough to eat but a low-paying job and a minor ailment: curing his ailment and drastically increasing his pay, perhaps? (To avoid confounding intuitions, it is important to make the morally motivated stranger better off than the subject in both cases.) However you answer this question, you must pick a change that would be equally *prudentially motivating* for the subject as the daily sandwich would be when he suffers from constant hunger. To report my own thoughts: when I arrive at a change that genuinely seems prudentially equivalent (which must be a very large change), it is hard to maintain that the change in the first case has more moral value than the change in the second. Indeed, if the morally motivated stranger defers completely to the subject's own reasons for improving his well-being in each case, then he has by definition equal moral reason to help him in each case.

The situation doesn't appear to be different if we add in a person who is better off than our subject in both cases. When I am deciding whether to help our subject or this better-off person by a given amount, then if I should give our subject a daily sandwich when he is constantly hungry (rather than help the better-off person by the given amount), I should also cure his minor ailment and drastically increase his pay when he is not food-insecure, or whatever the equivalent change is (rather than help the better-off person by the given amount).

What will make a difference, according to the relative prioritarian, is if there are *two* people, one who is food-insecure and one who has a low-paying job and a minor ailment, and I am choosing between giving the first person a daily sandwich and giving the second one the prudentially equivalent change. In this case, the relative prioritarian will say that we have to prioritize the first person – that helping him by a fixed amount contributes more to the total good.

So far, I claim, these verdicts aren't counterintuitive. If we are considering a single-person case, we don't have more reason to help a single badly-off person than he has to help himself, nor less reason to help a single better-off person

<sup>33</sup>I thank an anonymous referee for this example.

than he has to help himself. But if we are considering multi-person cases, we have more reason to help a badly-off person by a given amount *rather than* a better-off person by that same amount. At the very least, these claims stem from a coherent way to aggregate the good.

If one is not convinced, the relative prioritarian may still have a leg to stand on. Recall that relative prioritarianism makes two claims: that moral value should be prudential value, and that the worse-off should get more moral weight. My particular concern has been to show that these two claims together provide a rival way to satisfy the prioritarian commitments, and that they constitute an independently compelling philosophical view. But there is actually a third alternative, which is to agree with the absolute prioritarian that moral value diminishes marginally in prudential value *and* to agree with the relative prioritarian that the worse-off get more moral weight; call this position *hybrid prioritarianism*. So if one is on board with giving the worse-off more weight, but finds the above verdicts counterintuitive on the grounds that we have stronger moral reasons to help the absolutely worse-off subject, then one could adopt hybrid prioritarianism.

It should be obvious how to define such a position: instead of taking a rank-weighted average of utility values (as in relative prioritarianism), take a rank-weighted average of moral values, where the latter is a concave function of utility (as in absolute prioritarianism).<sup>34</sup>

Similarly, one can combine relative prioritarianism with a position such as *sufficientarianism*, the view that certain needs get lexical priority – for example, that if some people are below the poverty line, then it contributes nothing to the good to improve the lot of people who are above the poverty line. Sufficientarianism itself is incomplete, since one still needs to say, within the group of individuals whose needs get lexical priority, how an improvement for each contributes to the total good: is it better to improve the lot of worse-off individuals ‘below the line’ by some utility value or to improve the lot of better-off individuals ‘below the line’ by some different utility value?<sup>35</sup> Again, this is a substantive question, and relative prioritarianism may provide an attractive answer. In short: taking into account the kinds of ‘absolute’ considerations that other views are interested in needn’t preclude *also* taking into account considerations of relative priority.

## 8. Whose good should we bring about?

I already mentioned a key feature that distinguishes relative from absolute prioritarianism: relative prioritarianism denies strong separability, i.e. relative prioritarianism holds that whether the overall good is higher when Ann and Bob

<sup>34</sup>Now that we’ve seen that the weighting, assessment and attribute questions are independent, it should also be obvious how to define a number of other views, e.g. an absolute prioritarianism that also cares about inequality.

<sup>35</sup>See Crisp (2003) for one answer to this question. As far as I can interpret his view in the framework of this paper, he advocates a type of absolute prioritarianism below the sufficiency line, and utilitarianism above it.

have some utility amounts or when they have some different utility amounts depends on what Cecil has. To make it concrete: {ANN, 125; BOB, 200; CECIL, 150} may be better than {ANN, 100; BOB, 300; CECIL, 150}, but {ANN, 100; BOB, 300; CECIL, 0} may be better than {ANN, 125; BOB, 200; CECIL, 0}, even though the two comparisons involve the same utilities for Ann and Bob. In the first comparison, Ann's welfare makes a much higher contribution to overall good than Bob's, because she is the worst-off and Bob is the best-off; but in the second comparison Ann's welfare makes an only somewhat higher contribution, because she is the middle and Bob is the best-off. In this section, I will consider whether this should be thought of as a reason to favour absolute prioritarianism, and I will close with some remarks about how the good relates to what to do.

I've already explained why the relative prioritarian rejects strong separability: if Ann is worse-off than Bob, then whether she is the worst-off or the middle will make a difference to the relative contribution of Ann's and Bob's welfare to the total good, and Cecil's welfare determines where Ann and Bob fall in the relative ordering. So there is nothing inconsistent about denying strong separability and still holding that individual welfare is the only thing of value – that relational properties aren't themselves valuable. Nonetheless, let me raise a challenge to the denial of strong separability.

The challenge is a version of Parfit's 'divided world' example (Parfit 1991: 87–88, 99–100). Imagine the world contains only Ann, Bob and Cecil, but that Ann and Bob do not know of Cecil's existence, and vice versa. Whether Cecil is at 150 or 0 will make a difference to whether it is better for Ann to be at 125 and Bob 200 (option 1), or for Ann to be at 100 and Bob 300 (option 2). But how could this make a difference, if Cecil is wholly detached from Ann and Bob?

To see how the relative prioritarian would reply, notice first that overall good is the good of a *particular set of people* – overall good is indexed to the people in a distribution. Thus, there are two ways in which one or the other option could be better: it could be better for the group consisting of Ann and Bob, or it could be better for the group consisting of Ann, Bob and Cecil.

According to relative prioritarianism, there will be a single answer to the question of whether option 1 or option 2 is better – contains more overall good – for the group consisting of Ann and Bob. Cecil's well-being does not make a difference to which option is better for *Ann and Bob*, as it should be. (One view about *what to do* is to hold that we should decide between alternatives by considering which alternative is relative-prioritarian better for the group consisting of only the people involved, e.g. that we should do what is better for Ann and Bob. As it turns out, this is equivalent to a view known as the *competing claims view*.<sup>36</sup>)

Cecil's well-being *will* make a difference to which option is better for *Ann and Bob and Cecil*, that is, for the group consisting of all three. But I claim that there is nothing amiss here either, because Cecil is himself part of this group.

<sup>36</sup>This view holds that "we decide between alternatives by considering the comparative strength of the claims of different individuals, where (i) a claim can be made on an individual's behalf only if his interests are at stake; and (ii) his claim to have a given alternative chosen is stronger: (iia) the more his interests are promoted by that alternative and (iib) the worse off he is relative to others with whom his interests conflict" (Voorhoeve and Fleurbaey 2012: 397).

One might worry: how can Cecil's well-being make a difference to which option is better for all three, when Cecil fares the same under each option? The answer is that his well-being makes a difference to the overall effect of Ann's and Bob's well-being on the whole. If we take seriously the idea that the overall good (of a group of people) is more strongly related to the well-being of its worse-off member than its best-off member, then it is unsurprising that changes in the well-being of one member can influence how the well-being of other members affects the overall good – a person whose well-being was at some point only weakly determinant of the overall good can become more strongly determinant of the overall good, in virtue of her now becoming the worst-off member.

The cost of relative prioritarianism, then, isn't really that it gives the wrong verdict in divided world cases, since if one likes the verdict that Cecil's well-being doesn't make a difference to the good when she is unaffected, we can get this verdict by specifying that we are talking about the good for Ann and Bob. The real cost – if there is one – is that good (and hence the ordering of options) will always be relativized to a set of people. Unlike views that accept strong separability, the relative prioritarian view might give a different answer about which of two things increases the good more, when we are talking about the good of a smaller group or the good of a larger group which contains that smaller group. Thus, when we are answering the question *what to do*, we have to ask *from the point of view of increasing the good of what set of people?*

But this is just a general question we face in other ways in practical ethics. One is part of a family, a neighbourhood, a country, and humanity, and sometimes the good for each of these groups conflicts: what's good for the members of my family might be bad for the citizens of my country, and what's good for the citizens of my country might be bad for humanity. (Even though, in this case, there are no conflicts between various parts of a group – between Ann and Bob on the one hand and Cecil on the other – there is still a conflict between what is good for the smaller group and what is good for a larger group because the individuals occupy different positions in those groups and thus contribute differentially to their overall good.) And thus when making a moral decision, we must figure out whose good is the one to pay attention to.

We are left, then, with an additional question in distributive ethics: not just 'what should I do to bring about the good?', but '*whose* good should I be concerned with bringing about?'. We could hold that we should bring about the good of those whose interests are at stake in this particular choice; or the good of our present society; or the good of present and future people anywhere; or the good of humanity as a whole. I leave this question for further discussion.

## 9. Conclusion

Absolute prioritarianism holds that an individual's utility depends only on how good things are for her; that relational properties have no intrinsic value; and that nonetheless the measure of goodness in a society is spread-averse. Absolute prioritarianism adheres to these three claims by holding that we maximize average or total moral value, where moral value comes apart from utility (and,

indeed, diminishes marginally in utility). It therefore holds that the good of those who are worse-off in an absolute sense matters more than the good of those who are better-off in an absolute sense.

If one is convinced by these three claims, however, there is another alternative: we can hold that overall utility is a *weighted* average of utility, weighted towards those who are worse-off in a relative sense. Surprisingly, we can hold that the relatively worse-off matter more while still holding that an individual's utility depends only on how good things are for her (it does not depend on relational properties) and that relational properties have no intrinsic value. In particular, we can hold that the overall good in a society is more sensitive to the good of the worse-off than to the good of the better-off.

There are thus three different reasons for holding that distributions in which utility is more spread out are worse, keeping average utility fixed: because inequality is bad in itself (egalitarianism), because differences in utility matter more the worse off an individual is in an absolute sense (absolute prioritarianism), or because the relatively worse off get more weight (relative prioritarianism). I think that relative prioritarianism captures how we should think about overall good; but whether or not I've convinced you that it is the correct ethical view, it should be clear that relative prioritarianism is a serious contender.

**Acknowledgements.** This work was greatly improved by comments from Marc Fleurbaey, Johann Frick, Liz Harman, Niko Kolodny, Sven Neth, Kian Mintz-Woo, Tom Parr, Philip Pettit, Stephen White, and three anonymous referees; and by discussions at Cal State LA, UC Berkeley, the Formal Epistemology Workshop, the Princeton Workshop in Normative Philosophy, the Princeton Workshop in Philosophy and Economics, Harvard University, the Princeton University Center for Human Values, the University of Pittsburgh, UNC Chapel Hill, Duke University, Wayne State University, and Rutgers University. This work was supported by a Laurance S. Rockefeller Fellowship from the Princeton University Center for Human Values.

## References

- Adler M.** 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. New York, NY: Oxford University Press.
- Adler M.** 2019. *Measuring Social Welfare: An Introduction*. New York, NY: Oxford University Press.
- Arneson R.J.** 2000. Luck egalitarianism and prioritarianism. *Ethics* **110**, 339–349.
- d'Aspremont C. and L. Gevers** 2002. Social welfare functionals and interpersonal comparability. In *Handbook of Social Choice and Welfare*, Vol. **1**, ed. K.J. Arrow, A.K. Sen and K. Suzumura. Amsterdam: Elsevier Science.
- Blackorby C. and D. Donaldson** 1978. Measures of relative equality and their meaning in terms of social welfare. *Journal of Economic Theory* **18**, 59–80.
- Broome J.** 1991. *Weighing Goods: Equality, Uncertainty, and Time*. Oxford: Blackwell Publishers.
- Buchak L.** 2013. *Risk and Rationality*. Oxford: Oxford University Press.
- Chew S.H., E. Karni and Z. Safra** 1987. Risk-aversion in the theory of expected utility with rank dependent probabilities. *Journal of Economic Theory* **42**, 370–381.
- Crisp R.** 2003. Equality, priority, and compassion. *Ethics* **113**, 745–763.
- Fleurbaey M.** 2015. Equality versus priority: how relevant is the distinction? *Economics and Philosophy* **31**, 203–217.
- Hirose I.** 2009. Reconsidering the value of equality. *Australasian Journal of Philosophy* **87**, 301–312.
- Hirose I.** 2015. *Egalitarianism*. Abingdon: Routledge.
- Holtug N.** 2010. *Persons, Interests, and Justice*. Oxford: Oxford University Press.

- Holtug N.** 2019. Prioritarianism: ex ante, ex post, or factulist criterion of rightness? *Journal of Political Philosophy* **27**, 207–228.
- Jensen K.K.** 2003. What is the difference between (moderate) egalitarianism and prioritarianism? *Economics and Philosophy* **19**, 89–109.
- McCarthy D.** 2006. Utilitarianism and prioritarianism I. *Economics and Philosophy* **22**, 335–363.
- McCarthy D.** 2008. Utilitarianism and prioritarianism II. *Economics and Philosophy* **24**, 1–33.
- McCarthy D.** 2013. Risk-free approaches to the priority view. *Erkenntnis* **78**, 421–449.
- McCarthy D.** 2015. Distributive equality. *Mind* **124**, 1045–1109.
- Nagel T.** 1991. Equality. In *Mortal Questions*, 106–27. Cambridge: Cambridge University Press.
- Otsuka M.** 2012. Prioritarianism and the separateness of persons. *Utilitas* **24**, 365–380.
- Otsuka M.** 2015. Prioritarianism and the measure of utility. *Journal of Political Philosophy* **23**, 1–22.
- Otsuka M. and A. Voorhoeve** 2009. Why it matters that some are worse off than others: an argument against the priority view. *Philosophy and Public Affairs* **37**, 171–199.
- Parfit D.** 1991. *Equality or Priority? Lindley Lecture Series*. Kansas: University of Kansas, Department of Philosophy.
- Parfit D.** 2012. Another defence of the priority view. *Utilitas* **24**, 399–440.
- Persson I.** 2001. Equality, priority and person-affecting value. *Ethical Theory and Moral Practice* **4**, 23–39.
- Persson I.** 2008. Why levelling down could be worse for prioritarianism than for egalitarianism. *Ethical Theory and Moral Practice* **11**, 295–303.
- Persson I.** 2011. Levelling down and welfare diffusion. *Ethical Theory and Moral Practice* **14**, 307–311.
- Persson I.** 2012. Prioritarianism and welfare reductions. *Journal of Applied Philosophy* **29**, 289–301.
- Porter T.** 2010. Prioritarianism and the levelling down objection. *Ethical Theory and Moral Practice* **14**, 197–206.
- Rabinowicz W.** 2002. Prioritarianism for prospects. *Utilitas* **14**, 2–21.
- Rawls J.** 1971. *A Theory of Justice*. Cambridge, MA: Belknap Press of Harvard University Press.
- Rothschild M. and J.E. Stiglitz** 1970. Increasing risk: I. A definition. *Journal of Economic Theory* **2**, 225–243.
- Scheffler S.** 2000. *The Rejection of Consequentialism*. Oxford: Oxford University Press.
- Temkin L.S.** 1993. *Inequality*. New York, NY: Oxford University Press.
- Temkin L.S.** 1996. A continuum argument for intransitivity. *Philosophy and Public Affairs* **25**, 175–210.
- Temkin L.S.** 2003a. Egalitarianism defended. *Ethics* **113**, 764–82.
- Temkin L.S.** 2003b. Equality, priority or what? *Economics and Philosophy* **19**, 61–87.
- Tungodden B.** 2003. The value of equality. *Economics and Philosophy* **19**, 1–44.
- Voorhoeve A. and M. Fleurbaey** 2012. Egalitarianism and the separateness of persons. *Utilitas* **24**, 381–398.
- Voorhoeve A.** 2014. How should we aggregate competing claims? *Ethics* **125**, 64–87.
- Weber M.** 2019. The persistence of the leveling down objection. *Erasmus Journal for Philosophy and Economics* **12**, 1–25.
- Weymark J.A.** 1981. Generalized Gini Inequality Indices. *Mathematical Social Sciences* **1**, 409–30.
- Williams A.** 2012. The priority view bites the dust? *Utilitas* **24**, 315–331.

## Appendix 1. Aggregation Rules

Let a *society* consist of  $n$  individuals, and let  $X$  be the set of consequences. Define:

A *population distribution*  $e = \{1, x_1; \dots; n, x_n\}$  maps individuals  $i$  to consequences  $x_i$ .

$u_i(x_i)$  maps consequences  $x_i$  to real numbers, and represents the utility that individual  $i$  gets from  $x_i$  (we will use  $u_i$  as shorthand when the consequence is clear).

Let individuals be grouped into (mutually exclusive and exhaustive) groups  $G_1, \dots, G_m$  such that every individual in group  $G_j$  receives a consequence with utility value  $u_{G_j}$ , and let  $p(G_j)$  map groups  $G_j$  to  $[0, 1]$ .  $p$  represents the proportion of the population that is in  $G_j$ .

A *proportion distribution*  $d = \{p(G_1), u_{G_1}; \dots; p(G_m), u_{G_m}\}$  maps each group of size  $p(G_j)$  to consequences with utility  $u_{G_j}$ .

A few things to note about this latter definition: (1) each individual in  $G_j$  needn't receive the same consequence, as long as each receives a consequence with the same utility value; (2) the groups needn't

be the same for different distributions; and (3) for each distribution, there will typically be many equivalent ways to group individuals – all of the rules here will give the same result for each grouping.

A *population aggregation rule* assigns a numerical value to each population distribution  $e$ , such that of two population distributions, the one with the greater numerical value is better.

A *proportion aggregation rule* assigns a numerical value to each proportion distribution  $d$ , such that of two proportion distributions, the one with the greater numerical value is better. When we are dealing with a fixed population, as we are in this paper, every population rule gives rise to a proportion rule that will produce an equivalent ranking. The rules below are stated in terms of their ‘equivalent’ population and proportion versions.

### 1. Utilitarianism

The population version of utilitarianism is given by:

$$U_{total}(e) = \sum_{i=1}^n u_i$$

The proportion version of utilitarianism is given by:

$$U_{average}(d) = \sum_{j=1}^m p(G_j)u_{G_j}$$

### 2. Egalitarianism

The population version of egalitarianism is given by:

$$E_{total}(e) = \sum_{i=1}^n u_i - Q_{total}(e)$$

where  $Q_{total}$  maps population distributions  $e$  to real numbers and is a measure of how spread out  $e$  is.

The proportion version of egalitarianism is given by:

$$E_{average}(d) = \sum_{j=1}^m p(G_j)u_{G_j} - Q_{average}(d)$$

where  $Q_{average}$  maps proportion distributions  $d$  to real numbers and is a measure of how spread out  $d$  is. (To make these rules equivalent for a fixed population, define a suitable  $Q_{average}$  for each  $Q_{total}$ .)

### 3. Prioritarianism (‘Absolute prioritarianism’)

The population version of prioritarianism is given by:<sup>37</sup>

$$P_{total}(e) = \sum_{i=1}^n v(u_i)$$

where  $v(u)$  maps real numbers (utility values) to real numbers (moral values).

The proportion version of prioritarianism is given by:

$$P_{average}(d) = \sum_{j=1}^m p(G_j)v(u_{G_j})$$

where, again,  $v(u)$  maps real numbers (utility values) to real numbers (moral values).

For (absolute) prioritarianism,  $v$  can be taken to be strictly concave, or weakly concave but not linear.

<sup>37</sup>See footnote 7 for references.

**4. Rank-weighted utilitarianism ('Relative prioritarianism')**

From a population distribution  $e$ , define an *ordered population distribution*  $e'$  that reorders individuals from worst-off according to that distribution (lowest utility value) to best-off according to that distribution (highest utility value):

$$e' = \{1, x_1; \dots; n, x_n\}, \text{ where } u_1(x_1) \leq \dots \leq u_n(x_n).$$

Note that  $e'$  needn't be unique, because there may be ties, but any  $e'$  derived from the same  $e$  will yield the same value for the rule below.

From a proportion distribution  $d$ , define an *ordered proportion distribution*  $d'$  that reorders groups from worst-off according to that distribution (lowest utility value) to best-off according to that distribution (highest utility value):

$$d' = \{p(G_1), u_{G_1}; \dots; p(G_m), u_{G_m}\}, \text{ where } u_{G_1} \leq \dots \leq u_{G_m}.$$

Again,  $d'$  needn't be unique, but any  $d'$  derived from the same  $d$  will yield the same value for the rules below.

The population version of rank-weighted utilitarianism is given by:

$$W_{total}(e') = \sum_{k=1}^n \lambda_k u_k$$

where  $\lambda_k$  is a mapping from positive integers to (non-negative) real numbers, and represents the weight that the  $k^{\text{th}}$ -worst individual gets (d'Aspremont and Gevers 2002: 471).

When these weights are non-negative and non-increasing, i.e.  $\lambda_1 > 0$  and  $\lambda_k \geq \lambda_{k+1} \geq 0$  for  $k < m$ , then  $W_{total}$  defines the *generalized Gini family* (d'Aspremont and Gevers (2002: 471)).<sup>38</sup>

The proportion version of rank-weighted utilitarianism is given by:

$$W_{average}(d') = \sum_{j=1}^m \left\{ \left[ I \left( \sum_{i=j}^m p(G_i) \right) - I \left( \sum_{i=j+1}^m p(G_i) \right) \right] u_j \right\}$$

where  $I(p)$  measures the 'importance' or 'weight' of the interests of the top  $p$ -portion of individuals. In the main text, I've used  $w(p(G_i))$  to stand in for the expression in square brackets.

The proportion version of rank-weighted utilitarianism is equivalently given by:

$$W_{average}(d') = \sum_{j=1}^m \left\{ I \left( \sum_{i=j}^m p(G_i) \right) [u_j - u_{j-1}] \right\}$$

where  $I(p)$  again measures the 'importance' or 'weight' of the interest of the top  $p$ -portion of individuals and  $u_0$  is defined to be 0 (Note that  $I(\sum_{i=j}^m p(G_i))$  is the importance of the portion of individuals in groups  $G_j$  or higher).

For relative prioritarianism,  $I$  can be taken to be strictly convex, or weakly convex but not linear.

For a fixed population, the proportion version of rank-weighted utilitarianism, with the constraint that  $I$  is weakly convex, gives the same ordering as the population version of rank-weighted utilitarianism with the generalized Gini constraints. (The same holds for strict convexity if the Gini inequalities are strict.) We can

see this by setting  $I\left(\frac{n-(k-1)}{n}\right) = \frac{\sum_{i=k}^n \lambda_i}{\sum_{i=1}^n \lambda_i}$ .

**Appendix 2. Aggregation and SPREAD AVERSION**

All of the definitions and results in this section pertain to *proportion distributions*.

We can define *spread aversion in utility* formally, in two equivalent ways.

Define a Pigou–Dalton transfer to be one that removes utility of size  $a$  from an individual (or a group of size  $p$ ) and adds utility of size  $a$  to a worse-off individual (group of size  $p$ ), such that the latter individual

<sup>38</sup>The generalized Gini family was first introduced by Weymark (1981).

(group) remains no better off after the transfer.<sup>39</sup> Then we are strictly spread-averse if we think that a Pigou–Dalton transfer always makes a distribution strictly better, and we are weakly spread-averse if we think that a Pigou–Dalton transfer always makes a distribution weakly better (as-good-as-or-strictly-better).

Define a mean-preserving spread of  $d$  to be a distribution with the same mean utility as  $d$  and which can be obtained by a series of steps which consist in taking some proportion from the centre of the distribution and adding it to each tail, while preserving its mean value.<sup>40</sup> Then, we are strictly spread-averse if we think a mean-preserving spread always makes a distribution strictly worse (or if we think a mean-preserving ‘contraction’ – the inverse of a spread – always makes a distribution strictly better), and we are weakly spread-averse if we think a mean-preserving spread always makes a distribution weakly better (a mean-preserving contraction always makes a distribution weakly worse).

These definitions are equivalent. It is easy to see that a Pigou–Dalton transfer is a mean-preserving contraction. It is also easy to see that each step in a mean-preserving contraction is equivalent to a series of Pigou–Dalton transfers, and a series of such steps is thus equivalent to a series of Pigou–Dalton transfers.

That our aggregation rule is *spread averse in utility* will be our formal interpretation of SPREAD AVERSION.

Facts:

- (1) Utilitarianism is weakly spread-averse (in the trivial sense: a Pigou–Dalton transfer or a mean-preserving spread always preserve utilitarian value).
- (2) Egalitarianism: As long as  $I$  is suitably defined,  $E$  can be spread-averse in either the weak or strict sense.
- (3) Absolute prioritarianism: If  $v$  is weakly concave,  $P$  is weakly spread-averse. If  $v$  is strictly concave,  $P$  is strictly spread-averse.<sup>41</sup>
- (4) Relative prioritarianism: If  $I$  is weakly convex,  $W$  is weakly spread-averse. If  $I$  is strictly convex,  $W$  is strictly spread-averse.<sup>42</sup>

<sup>39</sup>d’Aspremont and Gevers (2002: 506) give the definition for population distributions; I adapt it here for proportion distributions.

<sup>40</sup>See Rothschild and Stiglitz (1970). They define this for a lottery – a probability distribution of consequences – but since it is merely a formal definition, we can interpret probabilities as proportions and locations of consequences as people rather than states of the world, and use the formal results to discuss proportion distributions rather than lotteries.

<sup>41</sup>See Rothschild and Stiglitz (1970) and Buchak (2013: 63n16), both interpreted for social choice (as in footnote 30). The latter cites a result from Chew *et al.* (1987: 375, Corollary 2).

<sup>42</sup>See Buchak (2013: 63n16), who cites a result from Chew *et al.* (1987: 375, Corollary 2), interpreted for social choice (as in footnote 30).