

ARTICLE

Calibration dilemmas in the ethics of distribution

Jacob M. Nebel^{1*}  and H. Orri Stefánsson^{2*} 

¹University of Southern California, School of Philosophy, 3709 Trousdale Parkway, Los Angeles, CA 90089, USA and ²Stockholm University, Department of Philosophy, Universitetsvägen 10 D, SE-106 91, Stockholm, Sweden

*Corresponding authors. Email: jnebel@usc.edu, orri.stefansson@philosophy.su.se

(Received 02 July 2020; revised 16 May 2021; accepted 21 September 2021; first published online 09 February 2022)

Abstract

This paper presents a new kind of problem in the ethics of distribution. The problem takes the form of several ‘calibration dilemmas’, in which intuitively reasonable aversion to small-stakes inequalities requires leading theories of distribution to recommend intuitively unreasonable aversion to large-stakes inequalities. We first lay out a series of such dilemmas for prioritarian theories. We then consider a widely endorsed family of egalitarian views and show that they are subject to even more forceful calibration dilemmas than prioritarian theories. Finally, we show that our results challenge common utilitarian accounts of the badness of inequalities in resources.

Keywords: Equality; prioritarianism; egalitarianism; utilitarianism; calibration

1. Introduction

Leading contemporary theories in the ethics of distribution can be regarded as attempts to stake out a reasonable middle ground between two extremes.

On one extreme is the utilitarian principle of distribution, which evaluates distributions according to their total (or average) well-being. Many philosophers reject utilitarianism because it is insensitive to inequality in the distribution of well-being. Many of us believe, contrary to utilitarianism, that we ought to give some priority to those who are worse off than others.

On the other extreme is the leximin (short for ‘lexicographic maximin’) rule, inspired by Rawls’s (1971) difference principle. Leximin assigns absolute priority to the very worst off. It says to choose, between any two distributions, the one that is better for the worst-off person whose welfare differs between those distributions.

Whereas utilitarianism seems to assign too little (i.e. zero) priority to the worse off, leximin seems to assign too much. For example, consider a large population in which everyone is quite poorly off. Suppose that we can either benefit one member

© The Author(s), 2022. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

of this population by an arbitrarily small amount, or benefit any number of people better off than her by an arbitrarily large amount. It seems at least permissible to benefit the many. But leximin would require us to benefit the one, even if the many are quite poorly off too.

Most contemporary theorists in the ethics of distribution agree that leximin is implausibly extreme. Yet they have said surprisingly little about how much priority we should give, instead of absolute priority, to the worse off.¹ Similarly, they have given no precise guidance for how to, say, balance the interests of the very worst off against the interests of the second-worst off.

For example, prioritarists believe that benefiting a person matters more the worse off that person is. When introducing this view, Derek Parfit (1991: 20) says that prioritarianism

does not tell us how much priority we should give to those who are worse off. On this view, benefits to the worse off could be morally outweighed by sufficient benefits to the better off. To decide what would be sufficient, we must simply use our judgement.

Fair enough. It is hard to see how else a prioritarian could decide when benefits to the better off would outweigh benefits to the worse off without, as Parfit says, simply using her judgement.

Similar remarks apply to contemporary versions of egalitarianism. Egalitarians believe that it is bad for some to be worse off than others. But, as egalitarians are quick to clarify, they don't care *only* about inequality. Egalitarians don't judge a situation in which everyone is equally miserable to be better than a situation in which everyone is happy but unequally so. Many egalitarians care both about decreasing inequality and about increasing total well-being (see Barry 1989: 79; Temkin 2003; Persson 2006). But, Parfit (1991: 5) suggests,

if we give weight to both equality and utility [i.e. well-being], we have no principled way to assess their relative importance. To defend a particular decision, we can only claim that it seems right.

Again, fair enough. It is hard to see how else an egalitarian could make decisions that involve tradeoffs between equality and total well-being, without simply using her judgement.

This does not mean, however, that prioritarists and egalitarians can consistently endorse any combination of distributive judgements about all cases. On theories of both kinds, our judgements about some tradeoffs commit us to judgements about other tradeoffs. These commitments, we argue, pose a problem for contemporary theories of distributive ethics that are supposed to give some priority, but not extreme priority, to those who are worse off than others.

Our argument proceeds through a series of 'calibration dilemmas'. Our initial dilemmas are based on Rabin's (2000) structurally similar calibration result for

¹Notable exceptions include Tungodden and Vallentyne (2005), Fleurbaey *et al.* (2009) and Pummer (2013: Ch. 3), whose results are in a similar spirit to ours.

expected utility theory. Rabin illustrated that seemingly reasonable aversion to risk when small sums of money are at stake commits an expected utility maximizer to seemingly unreasonable aversion to risk when larger sums of money are at stake. Similarly, in our calibration dilemmas, seemingly reasonable aversion to inequalities involving small differences in well-being commits prioritarianism and the most commonly defended version of egalitarianism to seemingly unreasonable aversion to inequalities involving larger differences in well-being – e.g. inequalities from which half the population would gain an arbitrarily large quantity of well-being. (We will soon see some examples.) These implications are not as extreme as leximin's. But, we believe, they are nonetheless implausible.

We start by characterizing a general class of views that includes prioritarianism, as typically understood, as a special case. We lay out a series of calibration dilemmas for these views. We then consider a widely endorsed family of egalitarian views and show that views in this family avoid the dilemmas for prioritarians. However, they are subject to no less forceful calibration dilemmas of their own, and other versions of egalitarianism fare no better. We then consider the implications of our calibration dilemmas for utilitarianism. Though we take our results to provide some support for the utilitarian's insensitivity to inequality in the distribution of well-being, they also pose a problem for prominent utilitarian explanations of the badness of inequalities in the distribution of resources (e.g. wealth inequality).

Those who do not want to interpret our calibration dilemmas as support for utilitarianism have a few options. For instance, prioritarians and egalitarians could respond to these dilemmas by giving up aversion to inequality when small quantities of well-being are at stake. An alternative response would be to simply accept the extreme implications of aversion to small-scale inequality. In the concluding section we consider these responses and suggest avenues for further research.

2. Prioritarian calibration dilemmas

2.1 Weak prioritarianism

Prioritarians believe that, other things being equal,² benefiting a person matters more, from a moral or social point of view, the worse off the person is. As formulated by Parfit (1991: 105), the thought is that 'utility [i.e. individual welfare] has diminishing marginal moral importance'.

Prioritarians often favour interventions that increase equality. In fact, given any fixed sum of welfare, the view always favours a more equal distribution of that sum. However, unlike standard versions of egalitarianism, prioritarianism arguably does not care about equality for its own sake. For, according to prioritarianism, the moral value of benefiting a person is completely determined by the person's own welfare, and depends in no way on how their welfare compares to others'.

To see how this works, we operationalize prioritarianism in the following standard way. Each individual's well-being has moral value. The moral value of a distribution can be represented as the sum of the moral value of each

²We assume throughout that things other than people's welfare are held fixed across all interventions and situations we consider. So, for instance, questions about desert or responsibility will not arise. We also assume a fixed population size.

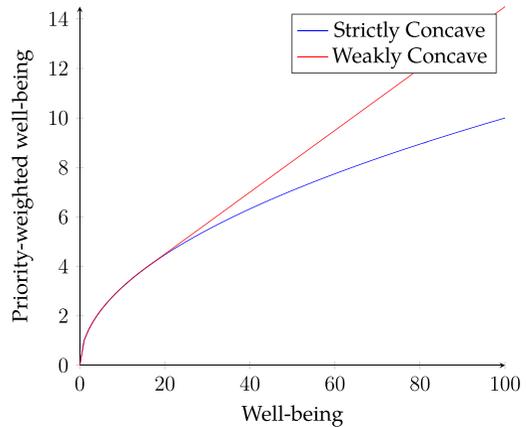


Figure 1. Strictly and Weakly Concave Priority Weighting Functions.

individual's well-being in that distribution.³ The moral value of a person's well-being does not depend on the existence or welfare of other people in the distribution. It is entirely a function of her own well-being. But the moral value of a person's well-being is not equal to her well-being; it is the value of a *priority weighting function* applied to her well-being. A person's priority-weighted well-being is a *strictly increasing, strictly concave* function of that person's well-being. To say that it is strictly increasing means that increasing a person's well-being will always increase her priority-weighted well-being. To say that it is strictly concave means that increasing a person's well-being by a fixed amount from a higher level increases her priority-weighted well-being by less than increasing her well-being by that same amount from a lower level. Together, these features mean that benefits to a person make a diminishing, but always positive, marginal contribution to the person's priority-weighted well-being.

A simple example of a strictly concave function is the square-root function, which is the blue curve depicted in Figure 1. The horizontal axis represents a person's well-being. The vertical axis represents her priority-weighted well-being. The slope of the square-root function is positive but decreases. That is what a strictly increasing, strictly concave function looks like.

Our problem applies to a much wider class of views than prioritarianism alone, so described. We will therefore treat the view just described as a special case, which we call *strict* prioritarianism. Strict prioritarians believe that the priority-weighting function is strictly concave, like the square-root function. *Weak* prioritarianism holds, more generally, that the priority-weighting function is *weakly* concave. This means that a fixed increase in well-being from a lower level increases a person's priority-weighted well-being by *at least as much* as a same-sized

³Although this framing is most natural for 'telic' prioritarianism, understood as a view about the *goodness* of distributions, our results apply just as well to deontic versions of prioritarianism (e.g. Williams 2012; Nebel 2017), as long as the choiceworthiness of a distribution, in the kinds of situations we discuss, is evaluated in a prioritarian manner. Our results are also robust across different views about the 'currency' of distribution. Although we explicitly discuss distributions of welfare, the same results apply for distributions of other goods – e.g. primary goods, capability indices, expected welfare, or whatever.

increase from any higher level. On this view, benefits need not have decreasing marginal value, but they cannot have increasing marginal value.

What is the interest of this more general class of views? Some people believe that, among individuals who are sufficiently well off, we should give no priority to those who are worse off (Crisp 2003). On one version of this *sufficientarian* view, benefits to those who are sufficiently well off still matter, but they should be distributed in a utilitarian manner. This view can be modelled as a version of weak prioritarianism in which the priority weighting function is strictly concave up to some point at which one has a sufficient quantity of well-being, after which it becomes linear. A simple example of such a function is the red curve in Figure 1; the red priority-weighting function follows the square-root graph until welfare level 16, at which point it becomes linear. Utilitarianism is another special case of weak prioritarianism – namely, one where the priority-weighting function is linear throughout. This means that priority is never given to those who are worse off (nor to those who are better off).

To summarize, the weakly prioritarian family of social welfare (i.e. moral value) functions compares distributions according to their sums of priority-weighted well-being, where the priority weighting function is strictly increasing and weakly concave. In the next subsection, we raise a problem for all views within this family except utilitarianism.

In laying out this problem, we appeal freely to numerical representations of welfare. Weak prioritarianism presupposes that such representations are meaningful. We judge that some tradeoffs involving such quantities are reasonable, others unreasonable. In this we follow the methodology suggested by Parfit: using our judgement.

Some may nevertheless find themselves unable to make judgements about tradeoffs involving different quantities of welfare without further information about how we assume welfare to be measured – i.e. what exactly the lives at various welfare levels are like. There are many views about the measure of well-being, and any particular choice of measure would be highly controversial. Our own attitude is that the tradeoffs we find reasonable seem (to us) reasonable on any plausible measure of well-being, and that the tradeoffs we find unreasonable seem (to us) unreasonable on any plausible measure as well. But some may find it helpful, for illustrative purposes, to have a concrete measure on the table.

One common practice in both the philosophical literature and in applications – e.g. in health economics – is to suppose that a unit of well-being corresponds to some duration of life of some quality (see e.g. Parfit 2012; Nord and Johansen 2014; Otsuka 2017; Lipman and Attema 2019). For example, we might take a unit of well-being to correspond to a year of very happy life. Because this sort of measure is both widely used and fairly concrete, we invite the reader to apply it when interpreting our results involving welfare differences. In section 2.3, however, we offer especially troubling results that do not depend on an arbitrary choice of welfare unit (and which, to our knowledge, have no direct precedent in the decision-theoretic literature).

2.2 Calibration results for weak prioritarianism

Suppose that your preferences between distributions are governed by weak prioritarianism.⁴ And suppose that you can benefit either of two people: Ann or Bob. Ann is slightly worse off than Bob: Bob has welfare w , Ann $w - 0.9$. For example, using a longevity-based measure of well-being, this could mean that Ann lives nine-tenths of a year less than Bob, where their lives are otherwise of equal quality. We represent this initial distribution as $(w - 0.9, w)$. You can either benefit Ann so that the two are equally well off, (w, w) , or benefit Bob by a slightly greater amount so that Bob becomes even better off, $(w - 0.9, w + 1)$. Suppose that, for some value of w , you would provide the smaller benefit to Ann, resulting in the equal distribution. This choice seems to us reasonable because, although benefiting Bob would maximize total welfare, the difference in total welfare is very slight; it is only one-tenth of a unit. For example, suppose that $w = 50$. We might think it reasonable to prefer $(50, 50)$ to $(49.1, 51)$. We will represent this preference in the following way (where ' $x \succ y$ ' means that x is preferred to y):

$$(50, 50) \succ (49.1, 51) \tag{1}$$

What can we infer from this particular preference about your preferences regarding tradeoffs with larger stakes? By itself, not very much; we need to know more about your preferences. To see what we can infer given slightly more information, let us start with the simplest, but perhaps somewhat unrealistic case: suppose, for now, that you would prefer the equal distribution for *any* value of w . (This assumption will soon be relaxed; that is, later we will assume that whether you prefer the equal distribution or the unequal one may depend on the size of w .)

We represent this more general preference as follows:

$$\text{For all } w : (w, w) \succ (w - 0.9, w + 1) \tag{2}$$

This preference may still seem reasonable because, for any value of w , the difference in total welfare between the equal and unequal distributions remains very small (0.1); it seems reasonable to benefit the worse off rather than benefiting the better off by a slightly greater amount.

Knowing only that your preferences between distributions are weakly prioritarian, and that they satisfy (2), what can we infer about your preferences over tradeoffs with larger stakes? In particular, suppose that Ann is 8 units worse off than Bob ($w - 8, w$), and that you can either benefit Ann by 8 units (w, w) or provide some larger benefit of size G to Bob ($w - 8, w + G$). What is the largest value of G such that you would prefer (w, w) to $(w - 8, w + G)$ that we can infer from our limited knowledge of your distributive preferences?

It might be natural to guess 10, 50, or 100, or to think that we cannot know without knowing more about the shape of your priority weighting function. In fact, however, we can – surprisingly – infer that you would prefer to benefit Ann by 8 units rather than benefiting Bob by G units *no matter how large G is*.

⁴By this we mean that, for some weakly concave priority weighting function, you prefer x to y just in case x contains at least as much priority-weighted well-being as y .

The above implication holds not only for tradeoffs between two individuals; it holds for tradeoffs between *any* number of people. And the same is true of the other implications that we shall discuss. The weak prioritarian evaluates a distribution according to the sum of each person's priority-weighted well-being, which cannot depend on the welfare or existence of other people. (That is the sense in which the prioritarian cares only about each individual's absolute welfare level.) So, if a weakly prioritarian distributor prefers a distribution in which two people have welfare w to a distribution in which one person has $w - L$ and the other has $w + G$, then she must also prefer a distribution in which any number of people have welfare w to a distribution in which half that number of people have $w - L$ and the other half have $w + G$.

It will be helpful, especially later on, to have some symbols to represent these distributions involving populations of arbitrary size. Take any distribution $\mathbf{w} = (w_1, \dots, w_n)$, where w_i represents the welfare of person i . If n people have the same welfare w , we represent this distribution as \mathbf{w}_n : that is, \mathbf{w}_n contains n people at level w . For any distribution \mathbf{w} and quantity of welfare k , let $\mathbf{w} + k$ represent the distribution $(w_1 + k, \dots, w_n + k)$ in which k is added to each person's welfare in \mathbf{w} . For any distributions $\mathbf{w} = (w_1, \dots, w_n)$ and $\mathbf{u} = (u_1, \dots, u_m)$, let (\mathbf{w}, \mathbf{u}) represent the distribution $(w_1, \dots, w_n, u_1, \dots, u_m)$ that concatenates (or 'stacks') \mathbf{w} and \mathbf{u} together. What we have found is that a weakly prioritarian distributor whose preferences satisfy (2) must also satisfy

$$\text{For all } w, n, \text{ and } G : \mathbf{w}_{2n} \succ (\mathbf{w}_n - 8, \mathbf{w}_n + G) \quad (3)$$

In words: a distribution in which $2n$ people are at any level w is preferred to one in which n people are at $w - 8$ and n are at $w + G$, no matter how large G is. This is, we think, an implausibly extreme degree of priority to the worse off. Intuitively, if w and G are very large, it is not reasonable to prefer the equal distribution to the unequal one; after all, a life at level $w - 8$ can be as good as we like, and the unequal distribution's net gain in total welfare, $(G - 8)$ times n , can be as large as we like.

The above result is one instance of our *prioritarian calibration theorem*, which is stated formally and proved in the Appendix; it is based on Rabin's calibration theorem for expected utility theory, which we mentioned in the Introduction. The general lesson of our prioritarian calibration theorem is this. Suppose that, for any welfare level w , a weakly prioritarian distributor would prefer any distribution in which two people are at level w to a distribution in which one person has welfare $w + g$ (greater) and another has $w - l$ (less), where g is greater than l (and both are positive). Then we can show that, for any welfare level w , such a distributor must prefer any distribution in which everyone would be at level w to a distribution in which half that population would have welfare $w + G$ and the other half would have $w - L$, where G is greater than L (and both are greater than g and l). If L is large enough, we can show this to hold for any value of G , however large. Put schematically, the theorem says that weak prioritarianism entails conditionals of the following form:

$$\begin{aligned} \text{If, for all } w : (w, w) \succ (w - l, w + g), \\ \text{then, for all } w, n : \mathbf{w}_{2n} \succ (\mathbf{w}_n - L, \mathbf{w}_n + G). \end{aligned} \quad (4)$$

Table 1. Weakly prioritarian distributor who prefers (w, w) to $(w - l, w + g)$ for all w when $g = 1$ must, for all w, n , prefer w_{2n} to $(w_n - L, w_n + G)$ for L s and G s entered in table

Then, for all $w, n: w_{2n} >$	If, for all $w, (w, w) > (w - l, w + 1)$				
	$l = 0.99$	$l = 0.95$	$l = 0.9$	$l = 0.75$	$l = 0.5$
$(w_n - 2, w_n +$	2)	2)	2)	5)	∞)
$(w_n - 4, w_n +$	4)	5)	7)	∞)	∞)
$(w_n - 6, w_n +$	6)	8)	20)	∞)	∞)
$(w_n - 8, w_n +$	8)	13)	∞)	∞)	∞)
$(w_n - 10, w_n +$	11)	21)	∞)	∞)	∞)
$(w_n - 15, w_n +$	17)	∞)	∞)	∞)	∞)
$(w_n - 20, w_n +$	25)	∞)	∞)	∞)	∞)
$(w_n - 25, w_n +$	33)	∞)	∞)	∞)	∞)
$(w_n - 50, w_n +$	105)	∞)	∞)	∞)	∞)
$(w_n - 75, w_n +$	∞)	∞)	∞)	∞)	∞)

Table 1 illustrates several representative examples of these results, where $g = 1$.⁵ Consider a weakly prioritarian distributor who, for all w , prefers (w, w) to $(w - l, w + 1)$, for each column's value of l . Then such a distributor must, for any w and any n , prefer w_{2n} to $(w_n - L, w_n + G)$, where each L is what is subtracted in a cell in the leftmost column in Table 1 while the addends in the other cells are the highest (integer) values of G for each choice of l and L . If a cell contains ∞ , this means that G can be arbitrarily large. For example, Table 1 says that a weakly prioritarian distributor who prefers (w, w) to $(w - 0.75, w + 1)$ for all w must, for all w, n , and G , prefer w_{2n} to $(w_n - 4, w_n + G)$.⁶

As Table 1 illustrates, the results can be made less extreme by requiring a larger value of l , as in the columns towards the left. Shifting the value of l towards the left can be regarded as shifting our preferences involving small-stakes tradeoffs closer to utilitarianism. For example, suppose we know only that a weakly prioritarian distributor prefers (w, w) to $(w - 0.99, w + 1)$ for any w - i.e. she is willing to forgo a measly hundredth of a unit of net total welfare to benefit the worse off. Such a

⁵These results can be easily generalized. What matters is the *ratio* between g and l . If the values of g and l in Table 1 are multiplied by a common constant k , the corresponding values of G and L should be multiplied by k . Our theorem also allows us to derive similar implications involving groups of different sizes. For example, a weakly prioritarian distributor who prefers $(w, w) > (w - 0.5, w + 1)$ for any w must, for any w, n , and G , prefer w_{100n} to $(w_n - 8, w_{99n} + G)$.

⁶What if the prioritarian wishes to treat negative welfare levels differently than positive ones? If the distributor's preference for (w, w) in the antecedent is restricted to $w \geq 0$, then the consequent would have to be restricted to $w \geq L$. This may seem a significant limitation on our results, because there is a certain kind of prioritarian - one who accepts the *ratio-scale invariance* condition introduced in section 2.3 - who must treat negative welfare levels differently (Brown 2007). As we will soon see, however, this kind of prioritarian will face their own, especially troubling calibration dilemma, which will only require a single preference at *some* welfare level. (Thanks to a referee for raising this question.)

distributor can have sensible preferences involving many somewhat larger tradeoffs. For example, she can prefer $(w - 25, w + G)$ to (w, w) for any integer G greater than 33. But, as shown in the bottom row, she must, for any w , prefer w_{2n} to $(w_n - 75, w_n + G)$, for arbitrarily large G . And even this seems unreasonable for sufficiently high values of w , since then $w - 75$ remains an excellent life.

The intuition behind our prioritarian calibration theorem is relatively simple. A weak prioritarian's aversion to inequality must be determined solely by the shape of her priority weighting function. A robust preference for benefiting the worse off when small quantities of well-being are at stake means that the marginal priority-weighted value of well-being diminishes very quickly, so that arbitrarily large gains above any given well-being level are less valuable than merely modest gains below that level.

To see how this works more concretely, we will walk through an informal sketch of the proof, applied to a particular choice of l and g . Suppose that we prefer (w, w) to $(w - 0.5, w + 1)$ for any w . And suppose that our preferences are weakly prioritarian. This means that, for some weakly concave priority weighting function $f(\cdot)$, and for any w , the total priority-weighted well-being of (w, w) is greater than the total priority-weighted well-being of $(w - 0.5, w + 1)$:

$$f(w) + f(w) > f(w - 0.5) + f(w + 1) \quad (5)$$

We want to know the marginal priority-weighted value of an increment of well-being at various levels. Suppose that some person has welfare level 100. What is the marginal priority-weighted value of an additional unit of welfare from this level? Well, we prefer $(100, 100)$ to $(100 - 0.5, 100 + 1)$, so our priority weighting function must be such that the total priority-weighted well-being of $(100, 100)$ is greater than the total priority-weighted well-being of $(99.5, 101)$:

$$f(100) + f(100) > f(99.5) + f(101) \quad (6)$$

So the difference in priority-weighted value between 100 and 101 must be less than the difference in priority-weighted value between 99.5 and 100:

$$f(101) - f(100) < f(100) - f(99.5) \quad (7)$$

Next consider the marginal priority-weighted value of an additional unit from level 101. Again, we prefer $(101, 101)$ to $(101 - 0.5, 101 + 1)$, so the difference in priority-weighted value between 101 and 102 must be less than the difference in priority-weighted value between 100.5 and 101:

$$f(102) - f(101) < f(101) - f(100.5) \quad (8)$$

But what is the difference in priority-weighted value between 100.5 and 101? Since our priority weighting function is weakly concave, the marginal priority-weighted value of welfare between 100.5 and 101 cannot exceed the marginal priority-weighted value of welfare between 100 and 101, since the latter interval starts from a lower level:

$$\frac{f(101) - f(100.5)}{101 - 100.5} \leq \frac{f(101) - f(100)}{101 - 100} \quad (9)$$

This means that the difference in priority-weighted value between 100.5 and 101 cannot be more than *half* the difference in priority-weighted value between 101 and 100:

$$f(101) - f(100.5) \leq 1/2[f(101) - f(100)] \quad (10)$$

So, putting (8) and (10) together, the difference in priority-weighted value between 102 and 101 must be less than half the difference between 101 and 100:

$$f(102) - f(101) < 1/2[f(101) - f(100)] \quad (11)$$

Iterations of this reasoning imply that the difference in priority-weighted well-being between consecutive whole-numbered levels is less than half the difference between the preceding levels. So, for example, the priority-weighted value of a unit of well-being gained from level 149 must be less than $(1/2)^{49} = 1/562949953421312$ times the difference in priority-weighted value between levels 100 and 101, which must in turn, by (7), be less than the difference between 99.5 and 100. This steep decline in marginal priority-weighted value establishes a firm upper bound on the priority-weighted value of gains in well-being. Specifically, no matter how large G is, the total difference in priority-weighted well-being between $100 + G$ and 100 must be less than twice the difference in priority-weighted well-being between 100 and 99.5.⁷ And this is true not just for gains from level 100, but from any w whatsoever.

The results in Table 1 seem pretty extreme. We expect that many proponents of weak prioritarianism would agree, and would therefore claim that a distributor should not prefer to benefit the worse off by a slightly lower amount from *all* levels of well-being. If w is sufficiently high, perhaps the distributor should no longer prefer to benefit the slightly worse off at the expense of total well-being. After all, weak prioritarians give priority to the worse off because of how badly off they are, and therefore need not prioritize the worse off when the worse off are extremely well off.

This move would prevent the weakly prioritarian distributor from having to prioritize the worse-off at the expense of arbitrarily large gains to the moderately better off. But it is not, by itself, enough to avoid implausibly extreme degrees of priority in large-stakes cases.

To see this, consider Table 2. Table 2 is based on the same calibration theorem as Table 1, but assumes only that the weakly prioritarian distributor prioritizes the slightly worse off at all levels of well-being up to 100. The theorem implies, for example, that if the distributor prefers any distribution in which both Ann and Bob have well-being level w to one in which Ann has $w - 0.75$ and Bob has $w + 1$ for any value of w up to 100, then she must prefer any distribution in which everyone would have well-being level 75 to one in which half that population would have well-being level 55 and the other half would have well-being level 1.67 million.

Some of the implications shown in Table 2 may seem reasonable – for example, for $l = 0.99$. More extreme results, however, can be delivered by adjusting the other parameters. For example, if a weakly prioritarian distributor would prefer (w, w) to $(w - 0.99, w + 1)$ for all values of w up to 3000, then she must prefer 500_{2n} to $(400_n, 5.98 \times 10^{12}_n)$. Or, if she would prefer (w, w) to $(w - 0.099, w + 0.1)$ for

⁷This is because the series $\sum_{i=1}^{\infty} (1/2)^{i-1}$ converges to 2.

Table 2. Weakly prioritarian distributor who prefers (w, w) to $(w - l, w + g)$ for all $w \leq 100$ when $g = 1$ must prefer 75_{2n} to $(75_n - L, 75_n + G)$ for L s and G s entered in table

If, for all w up to 100, $(w, w) > (w - l, w + 1)$					
Then, when $w = 75$, for any n , $w_{2n} >$	$l = 0.99$	$l = 0.95$	$l = 0.9$	$l = 0.75$	$l = 0.5$
$(w_n - 2, w_n +$	2)	2)	2)	5)	1.34×10^8)
$(w_n - 4, w_n +$	4)	5)	7)	6197)	9.40×10^8)
$(w_n - 6, w_n +$	6)	8)	20)	19263)	4.16×10^9)
$(w_n - 8, w_n +$	8)	13)	80)	42491)	1.70×10^{10})
$(w_n - 10, w_n +$	11)	21)	155)	83786)	6.86×10^{10})
$(w_n - 15, w_n +$	17)	56)	432)	387147)	2.20×10^{12})
$(w_n - 20, w_n +$	25)	101)	902)	1.67×10^6)	7.04×10^{13})
$(w_n - 25, w_n +$	33)	160)	1696)	7.05×10^6)	2.25×10^{15})
$(w_n - 50, w_n +$	80)	837)	26784)	9.39×10^9)	7.56×10^{22})
$(w_n - 75, w_n +$	141)	3728)	376242)	1.25×10^{13})	2.54×10^{30})

all values of w up to 100, then she must prefer 75_{2n} to $(25_n, 185418_n)$. And these results, based on seemingly reasonable degrees of priority to the worse off, still seem very extreme.⁸

In sum, prioritarrians face a dilemma: they can give seemingly reasonable priority to the worse off when small differences in well-being are at stake only by giving extreme – and, we think, unreasonable – priority to the worse off when very large differences are at stake.

2.3 Revenge dilemma for ratio-scale invariant prioritarianism

We do not insist on any particular response to this dilemma, but our sympathies lie in the following direction. If we are justified in giving priority to the worse off – that is, if utilitarianism is false – then weak prioritarianism is false. There is, however, room for other responses, corresponding to the different horns of our dilemma. Some weak prioritarrians may respond that, though we ought to give some priority to the worse off, it is unreasonable to prioritize the slightly worse off in the specific ways we have considered. Others may respond that it is reasonable to prioritize the moderately worse off even at the expense of the very large gains we have considered. We don't have anything like a decisive, non-question-begging argument against these

⁸An alternative response, suggested by an anonymous referee, would be to claim that the distributor should prefer the more unequal distribution when everyone is very *poorly* off, on the grounds that it is nearly just as urgent to benefit those who have slightly greater, but still very low, well-being. Our prioritarian calibration theorem applies to this sort of view as well, but we omit a detailed examination of its consequences because, as far as we know, there is no precedent for it in the literature. It does not seem very prioritarian to give *less* priority to the worse off at lower levels than at higher levels. However, the results of Table 2 can be easily adapted to this view: if the distributor prefers (w, w) to $(w - l, w + g)$ for all w between 10 and 110, then when $w = 85$, she must prefer $(w_n - L, w_n + G)$ for the values in the table. If the interval is narrower than $[10, 110]$, the results will of course be less extreme.

responses. We acknowledge that staunch proponents of weak prioritarianism who do not share our intuitions, or who are sufficiently willing to revise their intuitions about cases, may remain unfazed. But we ourselves would prefer a distributive theory, if there is one, that can accommodate the preferences we have reported.

To us, rejecting the intuitions is a desperate last resort. But there is a kind of sceptical attitude towards these intuitions that we would like to acknowledge. We have been freely expressing intuitions about cases involving various quantities of well-being. Some of these quantities seem small, others large. One may reasonably wonder whether we can rely on our intuitions about distributions expressed in these terms. After all, on many views about the measurement of well-being, such numbers are arbitrary. Unlike, say, money, there is no absolute scale of well-being. Numbers below 1 may *sound* small, and numbers above 100 may *sound* large. But, the sceptic may argue, we should not trust intuitions that are sensitive to particular numerical representations of welfare.

This response may seem to sap our prioritarian dilemma of its force. But it also inspires a new variation of our dilemma, which we take to be even more forceful than our initial version.

On most views about the measurability of welfare, a quantitative representation of well-being is only unique up to certain kinds of transformations. This is how we understand the idea that welfare has no absolute scale. The different kinds of admissible transformations correspond to different kinds of scales of well-being, and some theories of distribution are compatible with certain kinds of scales and not others. The most common kind of scale assumed by prioritarians is a *ratio* scale. This is a scale on which ratios of welfare levels are meaningful, but differences are not. For example, if Ann is represented as having welfare 50 and Bob welfare 100, then Bob is twice as well off as Ann, and any accurate representation of their welfare levels must assign Bob a welfare level that is twice as large as Ann's. But it need not be 50 units greater than Ann's, because the absolute size of the difference is arbitrary.

In the literature on social welfare functions, this idea is generally implemented in the following way. For any distribution $\mathbf{w} = (w_1, \dots, w_n)$ and number k , let $k\mathbf{w}$ represent the distribution (kw_1, \dots, kw_n) in which each person's welfare in \mathbf{w} is multiplied by k . According to the requirement of *ratio-scale invariance*, for any distributions \mathbf{w} and \mathbf{v} , and any positive k , \mathbf{w} is preferred to \mathbf{v} iff $k\mathbf{w}$ is preferred to $k\mathbf{v}$. This principle captures the idea that the only significant properties of our well-being measure are captured by well-being ratios: absolute differences and levels are not meaningful or morally significant. This principle is satisfied by the Atkinson social welfare function, which is the most widely endorsed version of prioritarianism. It is, for instance, defended at length by Adler (2012).⁹

If weak prioritarians deny that absolute differences in welfare are meaningful, then perhaps they can credibly insist that our intuitions about trading off 'small' and 'large' quantities of welfare should not be trusted. We can, however, derive analogous calibration results stated in terms of ratios. For example, suppose that a weakly prioritarian distributor would prefer a distribution in which Ann and

⁹Atkinson prioritarians hold that the priority weighting function $f(\cdot)$ has the form $f(w) = (1 - \delta)^{-1} w^{1-\delta}$ for $\delta > 0$ (except when $\delta = 1$, in which case $f(w) = \log w$), where δ represents the degree of priority to the worse off. For a critique of ratio-scale invariance, see Nebel (2021).

Bob are both at *some* level w to a distribution in which Bob is 1% better off than w and Ann is 0.5% worse off than w . And suppose that the distributor's preferences are ratio-scale invariant, in the sense that we have just explained. Then we can show that, for *any* level w , the distributor must prefer a distribution in which everyone is at w to a distribution in which half that population is 2% worse off than w and the other half has any quantity of welfare whatsoever.

This example follows from our *ratio-scale* prioritarian calibration theorem, which is also stated and proved in the Appendix. In schematic form, the theorem says that, given ratio-scale invariance, weak prioritarianism entails conditionals of the following form:

$$\begin{aligned} &\text{If, for some } w : (w, w) \succ ([1 - l]w, [1 + g]w), \\ &\text{then, for all } w, n : \mathbf{w}_{2n} \succ ([1 - L]\mathbf{w}_n, [1 + G]\mathbf{w}_n). \end{aligned} \quad (12)$$

Some representative results of this theorem are presented in Table 3. In the examples provided, we suppose that $g = 1\%$ – i.e. the better-off person in the slightly unequal distribution is 1% better off than w . Again, if a cell has ∞ , this means that G can be arbitrarily large. For example, the table says that, given ratio-scale invariance, a weakly prioritarian distributor who prefers (w, w) to $(0.9905w, 1.01w)$ for some w – e.g. $(100, 100)$ to $(99.05, 101)$ – must, for any w , any n , and any G , prefer $(\mathbf{w}_n, \mathbf{w}_n)$ to $(0.80\mathbf{w}_n, [1 + G]\mathbf{w}_n)$.

In Table 3, the only pairs of L and G that seem remotely plausible are when $l = 0.99\%$ – i.e. the worse-off person in the slightly unequal distribution would be only 0.99% worse off than w . For this value of l (when $g = 1\%$), G approaches infinity only as L grows close to 1: approximately $1 - 1.12 \times 10^{-30}$. So perhaps the weak prioritarian should insist that, when $g = 1\%$, the small-stakes inequality-introducing tradeoff should only be rejected when l is as high as 0.99%. However, we do not find this to be a very comfortable position for the prioritarian. We think this, first, because it seems no less reasonable to reject the tradeoff for slightly lower values of l , such as 0.95%. And, second, because it seems unreasonable to reject, from extremely low initial levels, the large-stakes tradeoff when L is close to 1 and G is arbitrarily large, since multiplying such extremely low levels by a minuscule $1 - L$ makes such a small difference, in particular when compared to the difference made by multiplying that same level by an arbitrarily large $1 + G$. If Ann's and Bob's lives are initially barely positive, then multiplying Ann's welfare by a factor arbitrarily close to zero will still result in her life being barely positive, but multiplying Bob's by an arbitrarily large factor could make his life wonderful.¹⁰

¹⁰An Atkinson prioritarian might try to rationalize these extreme implications by postulating that the moral value of welfare is bounded. After all, the particular members of the Atkinson family that have the most extreme results discussed above imply that the moral value of welfare is bounded (we thank Matthew Adler for emphasizing this). But even if one has a principled explanation for why the moral value of welfare should be bounded (which we do not find very plausible), this does not make it seem reasonable to reject the particular large-stakes tradeoffs that we have identified (e.g. tradeoffs in which one person is 2% worse off than w and another is better off by an arbitrarily large factor). What is extreme is not really the willingness to give up an infinite gain, but rather the small (potential) loss that is needed to outweigh any arbitrarily large (potential) gain. Moreover, even if there were independent reasons for thinking that the moral value of welfare is bounded, we take it to be instructive to learn that, assuming an Atkinson function, moderate inequality aversion when stakes are small *implies* that the moral value of welfare is bounded.

Table 3. Given ratio-scale invariance, a weakly prioritarian distributor who prefers (w, w) to $([1 - l]w, [1 + g]w)$ for some w when $g = 1\%$ must, for any w and n , prefer $(\mathbf{w}_n, \mathbf{w}_n)$ to $([1 - L]\mathbf{w}_n, [1 + G]\mathbf{w}_n)$ for $1 - Ls$ and $1 + Gs$ entered in table

If, for some w , $(w, w) > ([1 - l]w, 1.01w)$					
Then, for all w, n : $\mathbf{w}_{2n} >$	$l = 0.99\%$	$l = 0.95\%$	$l = 0.9\%$	$l = 0.75\%$	$l = 0.5\%$
$(0.98 \mathbf{w}_n,$	$1.02 \mathbf{w}_n)$	$1.02 \mathbf{w}_n)$	$1.02 \mathbf{w}_n)$	$1.03 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.96 \mathbf{w}_n,$	$1.04 \mathbf{w}_n)$	$1.04 \mathbf{w}_n)$	$1.05 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.94 \mathbf{w}_n,$	$1.06 \mathbf{w}_n)$	$1.07 \mathbf{w}_n)$	$1.15 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.92 \mathbf{w}_n,$	$1.08 \mathbf{w}_n)$	$1.11 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.90 \mathbf{w}_n,$	$1.11 \mathbf{w}_n)$	$1.17 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.85 \mathbf{w}_n,$	$1.16 \mathbf{w}_n)$	$1.93 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.80 \mathbf{w}_n,$	$1.23 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.75 \mathbf{w}_n,$	$1.30 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.50 \mathbf{w}_n,$	$1.97 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$
$(0.25 \mathbf{w}_n,$	$4.02 \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$	$\infty \mathbf{w}_n)$

Some may wonder whether we can rely on our intuitions about percentages of well-being. Even if such percentages are not arbitrary, it might be wondered whether we have a sufficiently good grip on them for us to be confident that certain tradeoffs are or are not reasonable. More generally, the reliability of people’s intuitions about quantities of well-being could be questioned.

Recall, however, the quotations from Parfit at the outset of this paper. Contemporary theories of distribution, such as prioritarianism and egalitarianism, attempt to avoid giving absolute priority to the worst off. But their views do not determine a unique intermediate degree of priority between absolute priority (leximin) and no priority (utilitarianism). To defend their distributive preferences, proponents of such theories must simply use their judgement. But if, as the sceptical attitude maintains, we cannot trust our intuitions about the kinds of tradeoffs we have presented, then how is any distribution-sensitive theorist supposed to form a particular view, or to make distributive choices in practice? If the sceptical response is right, we do not see how any particular degree of priority to the worse off could be justified. And, if we have no reasonable basis for some particular degree of priority, this would seem to make such views fairly uninteresting from a practical perspective, given the need to make tradeoffs in the real world.

3. Egalitarian calibration dilemmas

3.1 The generalized Gini family

Suppose we wish to avoid both horns of our calibration dilemma – that is, to accommodate moderate aversion to small-stakes tradeoffs without unreasonable aversion to large-stakes tradeoffs – and therefore reject weak prioritarianism. What kind of theory do we need?

The core commitment of prioritarianism that may appear responsible for our calibration results is its *additively separable* form. Each individual's well-being makes some contribution to the value of the welfare distribution. This contribution does not depend on the existence or welfare of other people, but is instead determined entirely by the priority-weighting function applied to her well-being. The value of a distribution is then represented as the sum of each individual's contributive value – that is, her priority-weighted well-being. This feature guarantees that the prioritarian distributor's responses to various tradeoffs depends entirely on the concavity of the priority weighting function. It therefore allows us to infer, from the rejection of certain small-stakes tradeoffs, severe constraints on the priority weighting function's degree of concavity at various welfare levels. These severe constraints lead to implausible results regarding large-stakes tradeoffs.

In light of the above observations, one might hope that the solution to our calibration dilemmas is to be found in a nonseparable social welfare function, such as paradigmatically egalitarian theories entail.¹¹ Unlike prioritarians, egalitarians care about how well or badly off people are in comparison to others. In particular, according to egalitarianism, the moral value that a person's welfare contributes to the total moral value of a welfare distribution partly depends on how well or badly off others are, which means that the moral value (or social welfare) function cannot be additively separable over individual welfare.

Perhaps the most widely endorsed version of egalitarianism is the *generalized Gini* family (Weymark 1981). We will focus on this particular family of egalitarian views in order to show, first, how egalitarians can avoid the calibration dilemma for prioritarianism, and, second, that egalitarians within this family face their own calibration dilemmas too. The motivation for focusing on this version of egalitarianism will become apparent below. In section 3.3, however, we briefly consider egalitarian views that fall outside this particular family of views.

A generalized Gini function evaluates distributions according to their weighted sums of well-being. Suppose that there are n people. We rank the n people from best-off to worst-off. Each person's welfare is multiplied by a weight, which depends on her rank. These weights can be represented as positive real numbers. The value of a distribution is the sum of the individuals' rank-weighted well-being. According to what we will call *weak egalitarianism*, the sequence of weights is *nondecreasing*, in the sense that the weight assigned to a person's well-being cannot be less than the weight assigned to anyone better-off than her. Utilitarianism is the version of weak egalitarianism where each person (irrespective of her rank) gets the same weight. *Strict egalitarianism* is the version of this view where the sequence of weights is *increasing*, so that the weight assigned to a person's well-being must be greater than the weight assigned to anyone better off than her. This is a generalization of the standard Gini social welfare function, whose weights are 1, 3, 5, 7, ... – i.e. the sequence of odd numbers. For our purposes, there is nothing special about this eponymous

¹¹When it comes to the distinction between egalitarianism and prioritarianism, we follow a terminological convention that has become common in philosophy since Parfit (1991); see also Broome (2015), Otsuka and Voorhoeve (2018) and Adler and Holtug (2019).

member of the generalized Gini family, whose weights seem to reflect ‘arbitrary distributional judgements’ (Weymark 1981: 414). We only require, with Weymark, that the weights be nondecreasing.

We define weak egalitarianism in this way for convenience, because it is the only paradigmatically egalitarian view that we extensively discuss.¹² We do not mean to imply that the generalized Gini family is the only kind of view that deserves the name ‘egalitarianism’, and we do not claim that our conclusions generalize to all other views that go by the name. There is, however, good reason to focus on the generalized Gini family. It can be derived from some axioms that should seem attractive to many theorists with broadly egalitarian intuitions.¹³ It is compatible with a merely cardinal (as opposed to ratio) scale of well-being; it plays nicely with both positive and negative welfare levels; and it has been claimed to have plausible implications in variable-population and risky cases (Asheim and Zuber 2014, 2016).

Finally, and most importantly, the generalized Gini family seems especially relevant to the present discussion in light of recent work by Lara Buchak. Buchak (2017) defends the generalized Gini family by appealing to her structurally analogous theory of decision under uncertainty. Buchak (2013: sec 2.3) claims that her decision theory avoids the problem posed by Rabin’s calibration theorem – i.e. that it can accommodate reasonable aversion to small-stakes risks without implying absurd aversion to larger-stakes risks. We might therefore suspect that the generalized Gini family – what we are calling weak egalitarianism – can avoid the problem posed by our prioritarian calibration theorem in a structurally analogous way.¹⁴

And indeed it does. To see this, suppose that a weakly egalitarian distributor would prefer a distribution in which both Ann and Bob are equally well off, at any level w , to one in which Bob has welfare $w + 1$ and Ann has welfare $w - 0.5$. This implies that the weight assigned to the worst-off, in a two-person society, is more than twice as great as the weight assigned to the better off person. This commits the distributor to preferring, for any level w , a distribution in which two people have welfare w to one in which one person has welfare $w - L$ and the other has $w + 2L$. That seems reasonable, not extreme. In sum, our prioritarian calibration theorem has no unwelcome implications for weak egalitarianism.

3.2 Calibration results for the generalized Gini family

We can, however, introduce a different calibration dilemma for weak egalitarianism, which seems to us even more forceful than our prioritarian dilemma. This result has a somewhat different setup, which will require us to consider several distributions

¹²Buchak (2017), however, defends the generalized Gini family on what she calls *prioritarian* grounds. Since the issue of nomenclature is irrelevant to our argument, we have no interest in defending our characterization of prioritarianism as committed to additive separability (thereby excluding nonutilitarian members of the generalized Gini family).

¹³See e.g. d’Aspremont and Gevers (2002) and Fleurbaey (2010). For more specific members of the Gini family, see Donaldson and Weymark (1980) and Blackorby *et al.* (2005: Ch. 4).

¹⁴For independent critiques of Buchak’s (2017) argument for the generalized Gini family, see Nebel (2020) and Stefánsson (2021).

over a larger population. Suppose that there are n people in some population, where n is even. And suppose that our preferences between distributions are governed by weak egalitarianism. Suppose, at first, that we can only affect the welfare of the two worst-off people in this population. In one distribution, both of these people would be equally well off, at some level w . In another distribution, one would be slightly better off than the other – one would have welfare $w + 1$, the other $w - 0.9$. Suppose further that, in the latter distribution, these two people would still be at the very bottom of the welfare distribution. This just requires that the third-worst-off person's welfare is no lower than $w + 1$; otherwise, their welfare can be whatever we like. It seems to us reasonable to prefer the former distribution – that is, to prefer equality between the two worst-off members of the population to small-stakes inequality between them – at least, for *some* values of w and of the welfare of other people in the population. Surely our priority to the relatively worse off should not get outweighed by a measly net gain of one-tenth of a unit of total welfare.

What could we conclude from this preference for equality when the stakes are small, given weak egalitarianism? By itself, not very much: just that the weight assigned to the second-worst-off person must be no more than nine-tenths of the weight assigned to the worst-off person. But suppose that this preference is not just confined to tradeoffs between the two worst-off individuals. Consider an analogous choice involving the second- and third-worst-off members of the population. In one distribution, both would be equally well off, at some level w (it needn't be the same value as in the previous paragraph). In another, one would have welfare $w + 1$, the other $w - 0.9$, where (again) this would not affect the rank ordering of members of the population. This just requires that the fourth-worst-off person's welfare is no lower than $w + 1$, and that the worst-off person's welfare is no greater than $w - 0.9$; otherwise, their welfare can be whatever we like. Again, it seems to us reasonable to prefer equality to small-stakes inequality between these two people – at least, for *some* values of w and of the welfare of other people in the population.

Suppose that we would have this pattern of preference for equality to small-stakes inequality between every pair of adjacent individuals in the rank ordering of welfare levels – again, for some (possibly different for each pair) value of w and some distribution over the rest of the population. (Perhaps this preference for equality between pairs of individuals of *any* rank is unreasonable; we will address that issue in a moment.) And suppose that $n = 10000$ – that is, we are assuming a population a third of the size of tiny San Marino's (we would get even more extreme results by assuming a larger population). Then, for any welfare level w , we must prefer a distribution in which ten thousand people have welfare w to one in which five thousand people have welfare $w - 2$ and the other five thousand have $w + 1.23 \times 10^{229}$. That seems to us implausibly extreme, particularly for high values of w .

This result is an instance of our egalitarian calibration theorem, which is stated and proved in the Appendix. The general lesson is this. Suppose that, for every pair of adjacent positions in a population's rank ordering of welfare levels, we prefer some distribution in which the two people in those positions have some equal welfare level, (w, w) , to a distribution in which one is slightly better off than the other, $(w - l, w + g)$, where $g > l > 0$, and everyone else's welfare and the rank-ordering

of individuals are unaffected by the choice (i.e. no one else's welfare level is between $w - l$ and $w + g$). And suppose that our preferences between distributions obey weak egalitarianism. Then, if the population is large enough, we must prefer a distribution in which everyone is equally well off at any w to any distribution in which half the population is moderately worse off ($w - L$) and the other half is much better off ($w + G$). This is because the weights assigned to better-off people must diminish so quickly that large benefits to such people are given very little weight. Table 4 shows some representative values of this theorem when $g = 1$ and $n = 10000$.

The results shown in Table 4 are extremely implausible, even when $l = 0.99$. Egalitarians might reply that the initial pattern of aversion to small-stakes inequality is unreasonable. What might seem unreasonable about this pattern is that it prefers equality in small-stakes tradeoffs even for the very best off. Egalitarians might insist that the aversion to small-stakes inequality is reasonable only when that inequality would be between members of some worse-off subset of the population. Suppose, for example, that it is only reasonable to prefer equality to small-stakes inequality between members of the worst-off decile of the population. The distributor would then still have to prefer any distribution in which everyone is equally well off at any level w to a distribution in which the one decile of the population would have welfare $w - L$ and everyone else would have $w + G$, for some extremely large values of G . Table 5 shows some representative values when, again, $g = 1$ and the population contains ten thousand people.

Since many egalitarians care about the distribution of welfare over all people (on some views, all animals) who ever live, they care about a population that is much larger than ten thousand people. And our calibration results are even more extreme for larger populations. For example, if there are one billion people, then even if the preference for equality over small-stakes inequality (where $l/g = 0.9$) is confined to the worst-off ten thousand people, one must prefer a distribution in which everyone has equal welfare at any level w to one in which ten thousand people have welfare $w - 2$ and everyone else has welfare $w + 6.76 \times 10^{449}$. That is extreme.

Some egalitarians may defend this extreme implication on the grounds that, when G is very large, the unequal distributions are *vastly* unequal, and at some point such vast inequalities should outweigh the gains in total well-being. But this response would put most contemporary egalitarians in distributive ethics in a difficult position – particularly those who favour the generalized Gini family. This is because such theorists are opposed to *levelling down* – i.e. decreasing the welfare of the better-off, leaving others unaffected, for the sake of equality. Most contemporary egalitarians believe, and the generalized Gini family implies, that increasing the welfare of some people, leaving others unaffected, always makes things better overall, even if it makes things worse with respect to equality. (Such egalitarians are, in Parfit's taxonomy, 'moderate.')

This is so even when the resulting inequality is vast. But it seems strange to insist that greatly increasing the well-being of some while leaving others unaffected always makes things better, no matter how much doing so increases inequality, while accepting the extreme implications we have identified – i.e. that greatly increasing the well-being of some (in the case above, 999990000 people) while making others *very slightly* worse off (e.g. by 2 units) makes things worse, owing to the resulting inequality.

Table 4. Weakly egalitarian distributor who prefers, for some w , (w, w) to $(w - l, w + g)$ at each pair of adjacent positions in a population of 10000 must, for any w , prefer \mathbf{w}_{10000} to $(\mathbf{w}_{5000} - L, \mathbf{w}_{5000} + G)$, for L s and G s entered in table

If, for each pair of adjacent positions, there is some w such that $(w, w) \succ (w - l, w + 1)$					
Then, for any w , $\mathbf{w}_{10000} \succ$	$l = 0.99$	$l = 0.95$	$l = 0.9$	$l = 0.75$	$l = 0.5$
$(\mathbf{w}_{5000} - 2, \mathbf{w}_{5000} +$	$1.33 \times 10^{22})$	$4.82 \times 10^{111})$	$1.23 \times 10^{229})$	$9.88 \times 10^{624})$	$2.82 \times 10^{1505})$
$(\mathbf{w}_{5000} - 4, \mathbf{w}_{5000} +$	$2.67 \times 10^{22})$	$9.64 \times 10^{111})$	$2.45 \times 10^{229})$	$1.98 \times 10^{625})$	$5.65 \times 10^{1505})$
$(\mathbf{w}_{5000} - 6, \mathbf{w}_{5000} +$	$4.00 \times 10^{22})$	$1.45 \times 10^{112})$	$3.68 \times 10^{229})$	$2.96 \times 10^{625})$	$8.47 \times 10^{1505})$
$(\mathbf{w}_{5000} - 8, \mathbf{w}_{5000} +$	$5.33 \times 10^{22})$	$1.93 \times 10^{112})$	$4.90 \times 10^{229})$	$3.95 \times 10^{625})$	$1.13 \times 10^{1506})$
$(\mathbf{w}_{5000} - 10, \mathbf{w}_{5000} +$	$6.67 \times 10^{22})$	$2.41 \times 10^{112})$	$6.13 \times 10^{229})$	$4.94 \times 10^{625})$	$1.41 \times 10^{1506})$
$(\mathbf{w}_{5000} - 15, \mathbf{w}_{5000} +$	$1.00 \times 10^{23})$	$3.61 \times 10^{112})$	$9.19 \times 10^{229})$	$7.41 \times 10^{625})$	$2.12 \times 10^{1506})$
$(\mathbf{w}_{5000} - 20, \mathbf{w}_{5000} +$	$1.33 \times 10^{23})$	$4.82 \times 10^{112})$	$1.23 \times 10^{230})$	$9.88 \times 10^{625})$	$2.82 \times 10^{1506})$
$(\mathbf{w}_{5000} - 25, \mathbf{w}_{5000} +$	$1.67 \times 10^{23})$	$6.02 \times 10^{112})$	$1.53 \times 10^{230})$	$1.23 \times 10^{626})$	$3.53 \times 10^{1506})$
$(\mathbf{w}_{5000} - 50, \mathbf{w}_{5000} +$	$3.33 \times 10^{23})$	$1.20 \times 10^{113})$	$3.06 \times 10^{230})$	$2.47 \times 10^{626})$	$7.06 \times 10^{1506})$
$(\mathbf{w}_{5000} - 75, \mathbf{w}_{5000} +$	$5.00 \times 10^{23})$	$1.81 \times 10^{113})$	$4.60 \times 10^{230})$	$3.70 \times 10^{626})$	$1.06 \times 10^{1507})$

Table 5. Weakly egalitarian distributor who prefers, for some w , (w, w) to $(w - l, w + g)$ at each pair of adjacent positions in the worst-off decile of a population of 10000 must, for any w , prefer \mathbf{w}_{10000} to $(\mathbf{w}_{1000} - L, \mathbf{w}_{9000} + G)$, for L s and G s entered in table

If, for each pair of adjacent positions in the lowest decile, there is some w such that $(w, w) \succ (w - l, w + 1)$					
Then, for any w , $\mathbf{w}_{10000} \succ$	$l = 0.99$	$l = 0.95$	$l = 0.9$	$l = 0.75$	$l = 0.5$
$(\mathbf{w}_{1000} - 2, \mathbf{w}_{9000} +$	509)	7.98×10^{19})	1.14×10^{43})	5.79×10^{121})	2.38×10^{297})
$(\mathbf{w}_{1000} - 4, \mathbf{w}_{9000} +$	1019)	1.60×10^{20})	2.29×10^{43})	1.16×10^{122})	4.76×10^{297})
$(\mathbf{w}_{1000} - 6, \mathbf{w}_{9000} +$	1528)	2.39×10^{20})	3.43×10^{43})	1.74×10^{122})	7.14×10^{297})
$(\mathbf{w}_{1000} - 8, \mathbf{w}_{9000} +$	2038)	3.19×10^{20})	4.58×10^{43})	2.32×10^{122})	9.52×10^{297})
$(\mathbf{w}_{1000} - 10, \mathbf{w}_{9000} +$	2547)	3.99×10^{20})	5.72×10^{43})	2.89×10^{122})	1.19×10^{298})
$(\mathbf{w}_{1000} - 15, \mathbf{w}_{9000} +$	3821)	5.98×10^{20})	8.58×10^{43})	4.34×10^{122})	1.79×10^{298})
$(\mathbf{w}_{1000} - 20, \mathbf{w}_{9000} +$	5095)	7.98×10^{20})	1.14×10^{44})	5.79×10^{122})	2.38×10^{298})
$(\mathbf{w}_{1000} - 25, \mathbf{w}_{9000} +$	6369)	9.97×10^{20})	1.43×10^{44})	7.24×10^{122})	2.98×10^{298})
$(\mathbf{w}_{1000} - 50, \mathbf{w}_{9000} +$	12739)	1.99×10^{21})	2.86×10^{44})	1.45×10^{123})	5.95×10^{298})
$(\mathbf{w}_{1000} - 75, \mathbf{w}_{9000} +$	19109)	2.99×10^{21})	4.29×10^{44})	2.17×10^{123})	8.93×10^{298})

Other egalitarians may respond to our result by choosing weights on welfare levels that avoid the extreme horn of our dilemma – that is, weights that don't imply absurd inequality aversion for high-stakes – by giving up seemingly reasonable aversion to inequality for smaller stakes. For example, as mentioned in the previous subsection, the weights of the standard Gini social welfare function are the odd numbers in increasing order. It is easy to see that using these weights would avoid the extreme consequences in Table 4. But this is because, in our view, the standard Gini is not even moderately inequality averse when stakes are relatively small and the population is sufficiently large. For example, the ratio of the weights between the 10th and the 11th best-off is $19/21$, which is greater than $9/10$. So the eponymous member of the generalized Gini family would *not* prefer that the 10th and 11th best-off people (or any adjacent pairs of individuals worse off than them) both be at welfare level w rather than one of them being at $w - 0.9$ and the other at $w + 1$. Nor would it prefer that the 20th and the 21st best-off people (or adjacent pairs worse off than them) both be at w rather than one being at $w - 0.95$ and the other at $w + 1$, or that the 100th and the 101st best-off people (or adjacent pairs worse off than them) both be at w rather than one of them being at $w - 0.99$ and the other at $w + 1$. So, given the standard Gini weights, none of the antecedents of the conditionals we have considered in this section would be satisfied, when applied to a population of more than 100 people. This, we suspect, would seem problematic to many who have egalitarian intuitions. The standard Gini function prefers equality to small-stakes inequality between adjacent individuals only among the very *best* off. But it seems to us strange to care so much more about inequality between, say, the best- and second-best-off individuals than about inequality between worse-off individuals in adjacent ranks.¹⁵

And yet, as our results illustrate, alternative members of the generalized Gini family that do seem sufficiently inequality averse when stakes are small are unreasonably extreme in their inequality aversion when stakes are large – hence the dilemma. More generally, our calibration dilemma for weak egalitarians is that intuitively reasonable aversion to small-stakes inequalities between small numbers of people (e.g. individual pairs, in our cases), even when limited to a small worse-off subset of the population, leads to manifestly unreasonable aversion to large-stakes inequalities between large numbers of people. The analogy between this dilemma and the calibration dilemma for weak prioritarianism is, very roughly, this: when the prioritarian prefers equality to small-stakes inequality at various welfare levels, this requires the marginal value of incremental benefits to diminish so quickly that very high welfare levels are given very little weight; when the egalitarian prefers equality to small-stakes inequality between individuals at various rank-positions, this requires the weights assigned to better-off people to

¹⁵It might be objected that the standard Gini is so sensitive to inequalities between (e.g.) the two best-off individuals not out of concern for unfairness between *them*, but rather because such inequalities would or might increase the distance from the best off to everyone else, and thus exacerbate the egalitarian complaints of the rest of the population. But this does not seem to explain the standard Gini's verdicts, because the ratio between the weights that the standard Gini assigns to adjacent members of the best-off group does not even depend on whether there are others worse off than them. (This feature is not shared by all members of the generalized Gini family; see Bossert 1990.)

diminish so quickly that large numbers of better-off people are given very little weight.¹⁶

Our dilemma for weak egalitarianism seems to us even more forceful than our dilemmas for weak prioritarianism, because the aversion to large-stakes inequalities is so extreme even when l/g is so close to 1 and L is relatively small. Although the egalitarian dilemma requires small-stakes inequality aversion to take a slightly more complicated form than the small-stakes inequality aversion assumed in the prioritarian dilemma, the assumed inequality aversion in the egalitarian dilemma is still quite weak. For each pair of ranks, we need only assume that the egalitarian prefers to benefit the worse off for *some* distribution meeting minimal constraints – i.e. for some value of w and some assignment of welfare levels to others that preserves the rank-ordering.

Of course, we have not considered other theories, besides the generalized Gini family, that deserve the name ‘egalitarian’. So we cannot claim to have identified a problem that afflicts all such views. However, our calibration results suggest that aversion to small-stakes inequalities is not best explained by the diminishing marginal importance of well-being (as the prioritarian thinks) or the diminishing weight given to the well-being of the better-off (as the generalized Gini egalitarian thinks).

3.3 Radically nonseparable egalitarianism

We mentioned above that the feature of weak prioritarianism that seems most clearly responsible for its calibration dilemma is its additive separability. We then showed that a similar dilemma afflicts the generalized Gini family, which is the most widely endorsed family of nonseparable views. One lesson we might draw from this result is that separability is not to blame for the prioritarian calibration dilemma. Another possible response, which should be more appealing to some egalitarians, is to claim that the generalized Gini family is still *too* separable, in a sense that we now explain.

In our egalitarian calibration dilemma, we considered small-stakes tradeoffs that each affect only two people. As long as those tradeoffs do not affect the rank ordering of anyone else in the population, the generalized Gini family allows us to ignore everyone unaffected by each tradeoff. And that is what allows us to infer constraints on the relative weight of each rank from the rejection of each

¹⁶This dilemma does not, to our knowledge, have any direct analogue in the decision-theoretic literature. The probabilistic calibration paradox of Sadiraj (2014) (see also Cox *et al.*, 2013) uses different small-risk lotteries, involving only two or three outcomes. The closest parallel is that of Safra and Segal (2008: sec. 4). The distributive analogue of their small-stakes condition, ‘stochastic B3’, defined on an interval of welfare levels $[a, b]$, would be as follows: any distribution w in which every welfare level is in $[a, b]$ and is instantiated by an even number of people is preferred to one in which half the people from each welfare level in w are better off by g and the other half are worse off by l . Two important differences between this condition and ours are that (1) it requires *every* individual to be affected in the small-stakes tradeoff, whereas our condition only involves two affected individuals at a time, and (2) it implies rejection of the small-stakes tradeoff for *any* distribution whose welfare levels are in the chosen interval, whereas ours only requires the (non-rank-switching) small-stakes tradeoff to be rejected from some initial distribution at each pair of ranks. Indeed, their condition is sufficiently strong that it cannot be satisfied by a generalized Gini function if the interval $[a, b]$ is sufficiently wide (this follows from their Proposition 2).

small-stakes tradeoff. This is a kind of separability, although not the full-fledged separability assumed by prioritarianism. It has been called *comonotonic separability* (see e.g. Wakker 2010: Ch. 10). Roughly, the moral value of distributions is comonotonically separable just in case, between two distributions that share the same rank ordering of persons,¹⁷ which of those distributions is better cannot depend on the welfare of unaffected individuals, that is, those individuals who have the same welfare in both distributions.

For theorists with broadly egalitarian inclinations, the lesson of our calibration results might be that comonotonic separability is still too strong. Such theorists might claim that, in order to value equality while avoiding absurd consequences, we must reject separability in a more radical way than the generalized Gini family does. We should, according to such theorists, accept a *radically nonseparable* version of egalitarianism. Such a theory may seem independently attractive to those who believe that distributions are organic unities (e.g. Broad 1930: 252).

Rejecting comonotonic separability, however, is not enough to avoid our calibration dilemmas. To see this, consider a simple egalitarian view, due to Rescher (1966: 35), that violates comonotonic separability.¹⁸ This view cares about two things: average well-being and inequality, measured (somewhat crudely) by a distribution's standard deviation. The relative importance of the standard deviation is represented by a positive weight x . The view values a distribution according to the difference between average well-being and this x -weighted standard deviation.¹⁹

Suppose that our preferences between distributions are governed by the egalitarian view just described. Now compare two distributions. In one, Ann and Bob have welfare w , and Cat and Dan have welfare $w + 100$. In another, Ann has welfare $w + 1$, Bob has $w - 0.99$, and Cat and Dan are still at $w + 100$. Suppose that, for some value of w , we would prefer the first distribution. Then we must prefer any distribution in which all four people would be equally well off, at any level w , to a distribution in which two would have welfare $w + 1000000$ and the other two would have $w - 1$.²⁰

Of course, this example is quite limited. It does not show that any radically nonseparable egalitarian view will be subject to calibration dilemmas. But it shows that not all such views avoid them. Radical nonseparability is not, by itself, sufficient to avoid calibration dilemmas. The example also suggests a more general challenge, if only in broad strokes. Any plausible egalitarian view would value at least two things: increasing aggregate welfare and decreasing inequality. In the radically nonseparable case, inequality is measured in such a way that the impact of a change in someone's well-being is sensitive to everyone else's well-being even when rank-ordering is held fixed. But, on any plausible measure of inequality, the effect of a small-stakes tradeoff between two members of a larger,

¹⁷More precisely, if one person is better off than another in one distribution, then she isn't worse off than that person in the other distribution.

¹⁸Thanks to Kacper Kowalczyk for bringing Rescher's view to our attention.

¹⁹That is, Average $- x \times$ Standard Deviation. Rescher sets $x = 1/2$. The problem we raise also applies if the weighted standard deviation is subtracted from total rather than average well-being.

²⁰This is because rejection of the small-stakes tradeoff requires x to exceed about 0.88, and acceptance of the large-stakes tradeoff requires x to be less than about 0.87.

already unequal population may be minuscule. So, to reject such a tradeoff, the distributor must assign extreme value to equality. This leads her to reject large-stakes transfers that greatly increase both aggregate welfare and inequality. That is what happens in our simple example, in which inequality is measured crudely by standard deviation. But similar dilemmas can be designed for other, more plausible measures of inequality.²¹

We therefore suspect that aversion to small-stakes inequality is not best explained by a concern for inequality understood as a holistic feature of distributions. But we do not insist on this lesson. We have not demonstrated the impossibility of a plausible, radically nonseparable version of egalitarianism that can capture reasonable aversion to small-stakes inequality without licensing extreme degrees of inequality aversion when the stakes are larger. We leave the search for such a view as a project for further research.

4 Utilitarianism and inequality

We take our calibration dilemmas for prioritarianism and egalitarianism to weaken the appeal that these views have over the extreme opposites of utilitarianism on the one hand and leximin on the other hand. But since leximin is, in our view, clearly too extreme, we think that our discussion so far should be seen as providing some indirect evidence in favour of utilitarianism. More on this in the next, concluding section. But first, we explain, in this penultimate section, why our calibration dilemmas also have some troubling implications for utilitarianism.

Utilitarianism, of course, is consistent with some aversion to certain kinds of inequality. For instance, utilitarians typically claim to be averse to inequality in the distribution of goods, such as money, that are assumed to have decreasing marginal utility (e.g. Bentham 1843; Pigou 2013). Some utilitarians (e.g. Greene and Baron 2001) appear to believe that the diminishing marginal utility of resources best explains and justifies our aversion to inequality in the distribution of resources, and indeed explains away our nonutilitarian intuitions about the badness of inequality in welfare.

We take our calibration results to count strongly against this utilitarian account of inequality aversion. For one can simply re-interpret the numbers in our calibration results for prioritarianism, from section 2.2, as, say, dollar amounts rather than magnitudes of welfare. Therefore, if a utilitarian wants to accommodate general inequality aversion with respect to money by postulating that people's welfare functions are strictly concave over money, then the

²¹Consider, for example, the 'total pairwise difference' model considered by Arrhenius (2013), based on Temkin's (1993) 'additive principle of equality' and 'relative to all those better off view of complaints' (see also Rabinowicz 2008). On this view, inequality is measured by the sum of (absolute) differences for each distinct pair of individuals. Suppose that the value of a distribution is given by aggregate welfare minus x -weighted inequality, so measured, for some real number x . Suppose, using the example raised above, that we prefer the first distribution $(w, w, w + 100, w + 100)$ to the second $(w - 0.99, w + 1, w + 100, w + 100)$ for some w . Then, for any w , we must prefer two thousand people having welfare w to one thousand having welfare $w + 1000000$ and the other thousand having welfare $w - 1$. For another calibration-like dilemma for this sort of view, see Gustafsson (2020).

resulting view will only be moderately inequality averse when relatively little money is at stake if it is absurdly inequality averse when more money is at stake.

For example, suppose that a utilitarian distributor would, on account of the decreasing marginal utility of money, prefer a situation where each of Ann and Bob have an equal total wealth w , up to \$100000, to a situation where Ann's total wealth is $w + \$1000$ while Bob's is $w - \$500$. Then our calibration result for weak prioritarianism shows that the distributor would have to prefer a situation where each of Ann and Bob have a total wealth of \$75000 to a situation where Bob has a total wealth of \$67000 and Ann has a total wealth of \$17 trillion (i.e. about 227 million times her initial wealth of \$75000). Of course, such an extreme difference in wealth may have negative effects on total welfare that cannot be explained in terms of the decreasing marginal utility of money. Our point is only that inequality aversion with respect to money cannot be *completely explained* by the decreasing marginal utility of money. A utilitarian who justifies aversion to small-stakes wealth inequality by pointing to the decreasing marginal utility of money would have to accept extreme consequences for large stakes *even if* the inequality in question would have no effects on total welfare that are not due to the decreasing marginal utility of money. But surely that is implausible.²² Although the social, structural, or political effects of great differences in wealth might justify the latter preference, it really is hard to see how an \$8000 dollar wealth loss to one person, from an initial wealth of \$75000, could *by itself* so drastically affect their standard of living that their welfare loss would be greater than the welfare gain to someone who, from the same initial wealth, increases their wealth by a factor of more than two hundred million.

There are other ways in which utilitarians can try to accommodate aversion to resource inequality. For example, some utilitarians claim that inequality is bad for people (see e.g. Broome 1991). On this kind of view, there might be some good – e.g. some instrumentally good resource, or perhaps individual well-being before considering the badness of inequality for each person – that ought to be distributed in an egalitarian manner, at least when other ingredients of well-being are held fixed.

We take our calibration results to count strongly against that view, as well, at least if an egalitarian distribution means that the distribution satisfies some version of the egalitarian formulas we have considered (e.g. the generalized Gini family). One can re-interpret the numbers in our egalitarian calibration results, from section 3, as quantities of some good rather than amounts of all-things-considered welfare. Therefore, a utilitarian like the one under consideration could only be moderately averse to inequality with respect to resources (or welfare without the inequality factor) when little is at stake and few people are involved if she is extremely averse to inequality when more is at stake and more people are involved.

²²This implication seems implausible even if the welfare effect of wealth has some limit – that is, even if there is some level of wealth beyond which no gain in wealth will increase a person's welfare. For such limits to justify the preference for (\$75000, \$75000) over (\$67000, \$170000000000000), one would have to assume that the welfare increase from extra wealth is satiated somewhere not too far from a total wealth of \$75000 dollars. But Kahneman and Deaton (2010: 16491) find that 'the effects of income on individuals' life evaluations show no satiation, at least to an amount well over \$120000' *per year*. Assuming that 'individuals' life evaluations' correlate with or contribute to their welfare it seems therefore unlikely that the effects of *total wealth* on welfare could be satiated not far from the \$75000 dollar level.

Of course, there may be other ways for utilitarians to explain our intuitive aversion to small-stakes inequalities in the distribution of certain goods. For instance, it may well be the case that *some* small-scale inequalities are bad for aggregate welfare, not (only) because of the decreasing marginal utility of money, but because of some other psychological phenomena such as envy or aversion to perceived distributive injustice. We ourselves however suspect that *if* utilitarianism is the correct theory of distribution, then at least most small-stakes tradeoffs like those that we have considered (both in welfare and in other goods) should probably be accepted, even though it seems reasonable to reject them. Moreover, since we think that our calibration results do provide some support for utilitarianism, we are inclined to take these results to cast doubt on egalitarian judgements condemning such small-scale tradeoffs.

We are not claiming that a utilitarian ought to, say, maximize total wealth, thus giving *no* weight to wealth equality. Our suggestion, rather, is that *small-stakes* inequalities are not as bad as our intuitive aversion to them would suggest; this is compatible with there being strong utilitarian reason to worry about much larger inequalities in, say, wealth. At the very least, our arguments suggest that a utilitarian's aversion to such small inequalities cannot be explained by such simple mechanisms as the diminishing marginal utility of money.

Again, we do not insist on any particular response to our dilemmas. But, however we respond to them, it appears that some of our considered judgements about the ethics of distribution will have to go.

5 Conclusion

Let us summarize our results. We have considered a broad generalization of prioritarianism and the most widely endorsed family of egalitarian views. We have shown that these theories can be moderately averse to small-stakes tradeoffs, in seemingly reasonable ways, only if they are extremely averse to large-stakes tradeoffs, in seemingly unreasonable ways. Moreover, we have seen that the most common utilitarian justification of inequality aversion (with respect to resources) can only explain such aversion when relatively little is at stake by implying extreme inequality aversion when more is at stake.

This leaves us with a few options.

One option is to conclude that the theories we have considered are false, on the grounds that our moderate aversion to small-stakes tradeoffs is reasonable but that the theories' resulting aversion to large-stakes tradeoffs is unreasonable. The research program suggested by this option would be to find some other theory that can successfully navigate both horns of our dilemmas. Our *hopes* lie with this option, but we are somewhat pessimistic about it. Though there may be various theories that avoid the particular calibration dilemmas we have raised, such theories may face other kinds of calibration dilemmas (as we saw with radically nonseparable egalitarianism) or be implausible in other ways. So, while we hope that others will discover a superior alternative to the distributive theories we have considered, we cannot take the existence of such an alternative for granted.

Another option is to embrace the extreme implications of the views we have considered. Those who are attracted to this option may take our results to demonstrate that considerations of justice and priority to the worse off are more demanding than we may have thought. We ourselves are not attracted to this option. If one is comfortable with the demands of justice or priority being so extreme, why not go all the way – that is, to leximin? Of course, our extreme implications for large-stakes tradeoffs are not quite as extreme as leximin's. But the position of rejecting leximin as implausibly extreme while welcoming the extreme attitude towards large-stakes tradeoffs in our calibration dilemmas seems to us somewhat unstable. At the very least, this option would seem to significantly weaken the attraction of the theories we have considered, which are often pitched as moderate and reasonable in contrast to leximin.

A third option is to maintain that our aversion to small-stakes tradeoffs is simply unreasonable. This option is compatible with many possible members of the weakly prioritarian and weakly egalitarian families, including versions that assign strict priority to those who are worse off. But it pushes these families closer to their common member: the utilitarian principle of distribution. *If* the small-stakes tradeoffs we have considered are unreasonable, then we ourselves would be inclined to place our bets on utilitarianism about distribution. If it were reasonable to give (strict) priority to the worse off, then we would have thought it reasonable to prefer a distribution in which two people are equally well off to an unequal distribution that has one-tenth of a unit less of aggregate welfare. At the very least, if we cannot reasonably be averse to the kinds of small-stakes tradeoffs we have considered, this would significantly weaken the attraction (for us) of nonutilitarian versions of weak prioritarianism and egalitarianism.

One might defend either of the second or third options by offering some debunking account of the intuitions to which we have appealed. For instance, proponents of the second option might argue that the large-stakes implications we have drawn out involve quantities that are greater than those that we can intuitively grasp. Proponents of the third option could argue that our aversion to small-stakes inequalities results from an unreliable overgeneralization of our attitudes towards real-world inequalities involving large quantities of welfare. Although there might be some truth to either or both of the above debunking claims, they do not – even if they are true – undermine the main message of this paper. Our main aim has been to show that some common and seemingly plausible intuitions about the ethics of distribution are not mutually satisfiable. And that is true, of course, regardless of the possibility of debunking some of these intuitions. Indeed, our calibration results are, in some ways, even more practically important for the theorist who is sceptical of the moral intuitions that make our results seem unpalatable. As competing versions of prioritarianism and egalitarianism support different degrees of priority to the worse off, our results may, at the very least, help such a sceptic to choose between different distributive theories.

Acknowledgements. Thanks to Gustaf Arrhenius, Matthew Adler, Ralf Bader, John Broome, Krister Bykvist, Kara Dreher, Marc Fleurbaey, Jimmy Goodrich, Hilary Greaves, Ben Holguin, Nils Holtug, Adam Kern, Kacper Kowalczyk, Dillon Liu, Andreas Mogensen, Michael Otsuka, Theron Pummer,

Wlodek Rabinowicz, Dean Spears, Elliott Thornley, and Alex Voorhoeve. The paper has been presented in seminars at University of Copenhagen, Stockholm University, and Institute for Futures Studies, as well as at the Fourth PPE Society Meeting, at the Fifth Bolder Workshop in Formal Value theory, and as the Parfit Memorial Lecture 2021 at the Global Priorities Institute in Oxford. We are grateful to the audiences for their comments and suggestions. Stefánsson gratefully acknowledges funding from Riksbankens Jubileumsfond (Pro Futura Scientia XIII).

References

- Adler M.** 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.
- Adler M and N. Holtug** 2019. Prioritarianism: a response to critics. *Politics, Philosophy & Economics* **18**, 101–144.
- Arrhenius G.** 2013. Egalitarian concerns and population change. In *Inequalities in Health*, eds N. Eyal, S.A. Hurst, O.F. Norheim and D. Wikler, 74–92. Oxford: Oxford University Press.
- Asheim G.B. and S. Zuber** 2014. Escaping the repugnant conclusion: rank-discounted utilitarianism with variable population. *Theoretical Economics* **9**, 629–650.
- Asheim G.B. and S. Zuber** 2016. Evaluating intergenerational risks. *Journal of Mathematical Economics* **65**, 104–117.
- Barry B.** 1989. *Theories of Justice: A Treatise on Social Justice*. Berkeley, CA: University of California Press.
- Bentham J.** 1843. *The Works of Jeremy Bentham, Vol. 1 (Principles of Morals and Legislation, Fragment on Government, Civil Code, Penal Law)*, eds J. Browning and W. Tait. Edinburgh: William Tait.
- Blackorby C., W. Bossert and D. Donaldson** 2005. *Population Issues in Social Choice Theory, Welfare Economics, and Ethics*. Cambridge: Cambridge University Press.
- Bossert W.** 1990. An axiomatization of the single-series Gini. *Journal of Economic Theory* **50**, 82–92.
- Broad C.D.** 1930. *Five Types of Ethical Theory*. London: K. Paul, Trench, Trubner.
- Broome J.** 1991. *Weighing Goods: Equality, Uncertainty and Time*. Oxford: Wiley-Blackwell.
- Broome J.** 2015. Equality versus priority: a useful distinction. *Economics and Philosophy* **31**, 219–228.
- Brown C.** 2007. Prioritarianism for variable populations. *Philosophical Studies* **134**, 325–361.
- Buchak L.** 2013. *Risk and Rationality*. Oxford: Oxford University Press.
- Buchak L.** 2017. Taking risks behind the veil of ignorance. *Ethics* **127**, 610–644.
- Cox J.C., V. Sadiraj, B. Vogt and U. Dasgupta** 2013. Is there a plausible theory for decision under risk? A dual calibration critique. *Economic Theory* **54**, 305–333.
- Crisp R.** 2003. Equality, priority, and compassion. *Ethics* **113**, 745–763.
- D’Aspremont C. and L. Gevers** 2002. Social welfare functionals and interpersonal comparability. In *Handbook of Social Choice and Welfare*. Vol. **1**, 459–541. Amsterdam: Elsevier.
- Donaldson D. and J.A. Weymark** 1980. A single-parameter generalization of the Gini indices of inequality. *Journal of Economic Theory* **22**, 67–86.
- Fleurbaey M.** 2010. Assessing risky social situations. *Journal of Political Economy* **118**, 649–680.
- Fleurbaey M., B. Tungodden and P. Vallentyne** 2009. On the possibility of nonaggregative priority for the worst off. *Social Philosophy and Policy* **26**, 258–285.
- Greene J. and J. Baron** 2001. Intuitions about declining marginal utility. *Journal of Behavioral Decision Making* **14**, 243–255.
- Gustafsson J.E.** 2020. The levelling-down objection and the additive measure of the badness of inequality. *Economics & Philosophy* **36**, 401–406.
- Kahneman D. and A. Deaton** 2010. High income improves evaluation of life but not emotional well-being. *Proceedings of the National Academy of Sciences USA* **107**, 16489–16493.
- Lipman S.A. and A.E. Attema** 2019. Rabin’s Paradox for health outcomes. *Health Economics* **28**, 1064–1071.
- Nebel J.M.** 2017. Priority, not equality, for possible people. *Ethics* **127**, 896–911.
- Nebel J.M.** 2020. Rank-weighted utilitarianism and the veil of ignorance. *Ethics* **131**, 87–106.
- Nebel J.M.** 2021. Utils and shmutils. *Ethics* **131**, 571–599.

- Nord E. and R. Johansen** 2014. Concerns for severity in priority setting in health care: a review of trade-off data in preference studies and implications for societal willingness to pay for a QALY. *Health Policy* **116**, 281–288.
- Otsuka M.** 2017. How it makes a moral difference that one is worse off than one could have been. *Politics, Philosophy & Economics* **17**, 192–215.
- Otsuka M. and A. Voorhoeve** 2018. Equality versus priority. In *Oxford Handbook of Distributive Justice*, ed. S. Olsaretti, 65–85. Oxford: Oxford University Press.
- Parfit D.** 1991. *Equality or Priority. The Lindley Lecture*. University of Kansas, Department of Philosophy.
- Parfit D.** 2012. Another defence of the priority view. *Utilitas* **24**, 399–440.
- Persson I.** 2006. A defence of extreme egalitarianism. In *Egalitarianism: New Essays on the Nature and Value of Equality*, eds N. Holtug and K. Lippert-Rasmussen, 83–98. Oxford: Clarendon Press.
- Pigou A.C.** 2013. *The Economics of Welfare*. London: Palgrave Macmillan.
- Pummer T.** 2013. The ethics of distribution. PhD dissertation. San Diego, CA: University of California.
- Rabin M.** 2000. Risk aversion and expected-utility theory: a calibration theorem. *Econometrica* **68**, 1281–1292.
- Rabinowicz W.** 2008. The size of inequality and its badness some reflections around Temkin's inequality. *Theoria* **69**, 60–84.
- Rawls J.** 1971. *A Theory of Justice*. Cambridge, MA: Harvard University Press.
- Rescher N.** 1966. *Distributive Justice: A Constructive Critique of the Utilitarian Theory of Distribution*. Indianapolis, IN: Bobbs-Merrill.
- Sadiraj V.** 2014. Probabilistic risk attitudes and local risk aversion: a paradox. *Theory and Decision* **77**, 443–454.
- Safra Z. and U. Segal** 2008. Calibration results for non-expected utility theories. *Econometrica* **76**, 1143–1166.
- Stefánsson H.O.** 2021. Ambiguity aversion behind the veil of ignorance. *Synthese* **198**, 6159–6182.
- Temkin L.S.** 1993. *Inequality*. New York, NY: Oxford University Press.
- Temkin L.S.** 2003. Egalitarianism defended. *Ethics* **113**, 764–782.
- Tungodden B. and P. Vallentyne** 2005. On the possibility of Paretian egalitarianism. *Journal of Philosophy* **102**, 126–154.
- Wakker P.P.** 2010. *Prospect Theory: For Risk and Ambiguity*. Cambridge: Cambridge University Press.
- Weymark J.A.** 1981. Generalized Gini Inequality Indices. *Mathematical Social Sciences* **1**, 409–430.
- Williams A.** 2012. The priority view bites the dust? *Utilitas* **24**, 315–331.

A. Appendix

Throughout the appendix, ‘ $\lfloor x \rfloor$ ’ denotes the greatest integer less than or equal to x , and ‘ $\lceil x \rceil$ ’ denotes the smallest integer greater than or equal to x .

Prioritarian Calibration Theorem. *Suppose that $f(\cdot)$ is strictly increasing and weakly concave over all welfare levels. Suppose that there are some welfare levels $\bar{w} > \underline{w}$ and $g > l > 0$ (where $g \leq 2l$) such that, for all $w \in [\underline{w}, \bar{w}]$, $2f(w) > f(w + g) + f(w - l)$. Then, for all $w \in [\underline{w}, \bar{w}]$, $L > 0$, and $G > 0$,*

(i)

$$f(w) - f(w - L) > \begin{cases} \sum_{i=1}^{\lfloor L/g \rfloor} (g/l)^i [f(w) - f(w - l)], & \text{if } L \leq w - \underline{w} \\ \left[(L - \lfloor w - \underline{w} \rfloor)(g/l)^{\lfloor (w - \underline{w})/g \rfloor} + \sum_{i=1}^{\lfloor (w - \underline{w})/g \rfloor} (g/l)^i \right] [f(w) - f(w - l)], & \text{if } L > w - \underline{w} \end{cases}$$

(ii)

$$f(w + G) - f(w) < \begin{cases} \sum_{i=1}^{\lceil G/g \rceil} (l/g)^{i-1} [f(w) - f(w - l)], & \text{if } G \geq \bar{w} - w \\ \left[(G - \lceil \bar{w} - w \rceil)(l/g)^{\lceil (\bar{w} - w)/g \rceil - 1} + \sum_{i=1}^{\lceil (\bar{w} - w)/g \rceil} (l/g)^{i-1} \right] [f(w) - f(w - l)], & \text{if } G < \bar{w} - w \end{cases}$$

Part (i) of our theorem is slightly stronger than Rabin’s – he uses increments of $2l$ rather than of g – so we include below the modified (and slightly simpler) proof of part (i). For proof of part (ii), see Rabin (2000).²³

Proof. $f(\cdot)$ is weakly concave, and $w - g < w - l$. So the marginal priority-weighted value of each unit of well-being over the interval $[w - g, w]$ – graphically, the slope of the priority weighting function over that interval – must be at least as great as it is over the interval $[w - l, w]$:

$$\frac{f(w) - f(w - g)}{g} \geq \frac{f(w) - f(w - l)}{l}.$$

Thus, $f(w) - f(w - g) \geq (g/l) [f(w) - f(w - l)]$. More generally, for any natural number k ,

$$f(w - [k - 1]g) - f(w - kg) \geq (g/l) [f(w - [k - 1]g) - f(w - [k - 1]g - l)].$$

If $w - g \geq \underline{w}$, then the tradeoff would be rejected from initial level $w - g$. So, $f(w - g) - f(w - g - l) > f(w) - f(w - g)$. And we have, from the paragraph above (letting $k = 2$), $f(w - g) - f(w - 2g) \geq (g/l)[f(w - g) - f(w - g - l)]$. So, $f(w - g) - f(w - 2g) > (g/l)[f(w) - f(w - g)]$. More generally, for any k such that $w - (k - 1)g \geq \underline{w}$,

$$f(w - [k - 1]g) - f(w - [k - 1]g - l) > f(w - [k - 2]g) - f(w - [k - 1]g).$$

Therefore, for any such k ,

$$f(w - [k - 1]g) - f(w - kg) > (g/l)[f(w - [k - 2]g) - f(w - [k - 1]g)].$$

So $f(w) - f(w - kg) > \sum_{i=1}^k (g/l)^i [f(w) - f(w - l)]$. This yields the lower bound on $f(w) - f(w - L)$ stated in the theorem.

²³The reader who wishes to recover the results of Table 1 or derive their own when the antecedent is satisfied for all w , without using the cumbersome formulation stated in the theorem, can use the following approximation of G in terms of l , g , and L :

$$G \approx \frac{g \log(2 - (\frac{g}{l})^{L/g})}{\log(\frac{l}{g})},$$

where G approaches ∞ as L approaches $\frac{g \log(2)}{\log(\frac{g}{l})}$.

Ratio-Scale Prioritarian Calibration Theorem. Suppose that $f(\cdot)$ is strictly increasing and weakly concave over all welfare levels, and satisfies ratio-scale invariance. Suppose there is some welfare level w and positive real numbers $\hat{l} < 1$ and $\hat{g} > 1/l$ such that $2f(w) > f(\hat{g}w) + f(\hat{l}w)$. Then, for all w ,

(i)

$$f(w) - f(\hat{l}w) > \sum_{i=1}^{\lfloor -\log(\hat{l})/\log(\hat{g}) \rfloor} \left(\frac{1 - \hat{g}}{\hat{g}(\hat{l} - 1)} \right)^i [f(w) - f(\hat{l}w)].$$

(ii)

$$f(\hat{G}w) - f(w) < \sum_{i=1}^{\lceil \log \hat{G} / \log \hat{g} \rceil} \left(\frac{\hat{g} - \hat{g}\hat{l}}{\hat{g} - 1} \right)^{i-1} [f(w) - f(\hat{l}w)].$$

Proof. We first prove part (i). By hypothesis, $f(w) - f(\hat{l}w) > f(\hat{g}w) - f(w)$. So, by ratio-scale invariance, for any natural number k ,

$$f(w/\hat{g}^k) - f(\hat{l}w/\hat{g}^k) > f(w/\hat{g}^{k-1}) - f(w/\hat{g}^k).$$

And, by weak concavity,

$$\begin{aligned} f(w/\hat{g}^{k-1}) - f(w/\hat{g}^k) &\geq \frac{w/\hat{g}^{k-1} - w/\hat{g}^k}{w/\hat{g}^{k-1} - \hat{l}w/\hat{g}^{k-1}} [f(w/\hat{g}^{k-1}) - f(\hat{l}w/\hat{g}^{k-1})] \\ &= \frac{1 - \hat{g}}{\hat{g}(\hat{l} - 1)} [f(w/\hat{g}^{k-1}) - f(\hat{l}w/\hat{g}^{k-1})]. \end{aligned}$$

So, for any k , $f(w/\hat{g}^{k-1}) - f(w/\hat{g}^k) > \frac{1 - \hat{g}}{\hat{g}(\hat{l} - 1)} [f(w/\hat{g}^{k-2}) - f(w/\hat{g}^{k-1})]$. Therefore,

$$f(w) - f(w/\hat{g}^k) > \sum_{i=1}^k \left(\frac{1 - \hat{g}}{\hat{g}(\hat{l} - 1)} \right)^i [f(w) - f(\hat{l}w)].$$

We next prove part (ii). By hypothesis, $f(\hat{g}w) - f(w) < f(w) - f(\hat{l}w)$. So, by ratio-scale invariance,

$$f(\hat{g}^k w) - f(\hat{g}^{k-1} w) < f(\hat{g}^{k-1} w) - f(\hat{l}\hat{g}^{k-1} w).$$

And, by concavity,

$$f(\hat{g}^k w) - f(\hat{l}\hat{g}^k w) \leq \frac{\hat{g}^k w - \hat{l}\hat{g}^k w}{\hat{g}^k w - \hat{g}^{k-1} w} [f(\hat{g}^k w) - f(\hat{g}^{k-1} w)] = \frac{\hat{g} - \hat{g}\hat{l}}{\hat{g} - 1} [f(\hat{g}^k w) - f(\hat{g}^{k-1} w)].$$

So, for any k , $f(\hat{g}^k w) - f(\hat{g}^{k-1} w) < \frac{\hat{g} - \hat{g}\hat{l}}{\hat{g} - 1} [f(\hat{g}^{k-1} w) - f(\hat{g}^{k-2} w)]$. Therefore,

$$f(\hat{g}^k w) - f(w) < \sum_{i=1}^k \left(\frac{\hat{g} - \hat{g}\hat{l}}{\hat{g} - 1} \right)^{i-1} [f(w) - f(\hat{l}w)].$$

Egalitarian Calibration Theorem. Let $I = \{1, \dots, n\}$ and $I_m = \{m, \dots, n - 1\}$. And let $w_{(j)}$ denote the welfare of the j th best-off individual. Suppose that for all adjacent pairs $j, j + 1 \in I_m$, there is some initial level $w = w_{(j)} = w_{(j+1)}$, where $w_{(j-1)} \geq w + g$ and $w_{(j+2)} \leq w - l$, such that $\sum_{i \in I} a_i w_i > \sum_{i \in I \setminus \{j, j+1\}} a_i w_i + a_j(w + g) + a_{j+1}(w - l)$. Then, for any w ,

(i)

$$\sum_{i=k}^n a_i w \geq \begin{cases} \sum_{i=1}^{n-k} (g/l)^{i-1} a_k w, & \text{if } k \geq m \\ \sum_{i=1}^k a_k w + \sum_{i=1}^{n-m} (g/l)^{i-1} a_k w & \text{if } k < m. \end{cases}$$

(ii)

$$\sum_{i=1}^k a_i w \leq \begin{cases} \sum_{i=1}^k (l/g)^{i-1} a_k w, & \text{if } m = 1 \\ \sum_{i=1}^m a_k w + \sum_{i=1}^{k-m} (l/g)^{i-1} a_k w, & \text{if } 1 < m < k \\ \sum_{i=1}^k a_k w, & \text{if } m \geq k \end{cases}$$

Proof. From $\sum_{i \in I} a_i w_i > \sum_{i \in I \setminus \{j, j+1\}} a_i w_i + a_j(w + g) + a_{j+1}(w - l)$, subtract $\sum_{i \in I \setminus \{j, j+1\}} a_i w_i$ from both sides. This yields $a_j w + a_{j+1} w > a_j(w + g) + a_{j+1}(w - l)$. Thus $a_{j+1} l > a_j g$, so for every $j \in I_m$, $a_{j+1} > (g/l)a_j$. Along with the nondecreasingness of the weights, this yields the upper and lower bounds on total rank-weighted utilities stated in the theorem.

Jacob M. Nebel is Assistant Professor of Philosophy at the University of Southern California. His research interests include the ethics of population, distribution, and risk; the structure of value relations; and the measurement and aggregation of well-being and other values.

H. Orri Stefánsson is Associate Professor of Practical Philosophy at Stockholm University, Pro Futura Scientia fellow at the Swedish Collegium for Advanced Study, and advisor at the Institute for Futures Studies. He currently works on decision-making under extreme uncertainty, population ethics, distributive ethics, and climate ethics.

Cite this article: Nebel JM and Stefánsson HO (2023). Calibration dilemmas in the ethics of distribution. *Economics & Philosophy* 39, 67–98. <https://doi.org/10.1017/S0266267121000298>