Efficient Inequalities

Jacob Barrett

1. Introduction

In his recent book, *Why Does Inequality Matter?*, T. M. Scanlon considers a number of objections to existing inequalities in the distribution of social and economic benefits. After discussing various negative consequences of such inequalities, he turns to the question of whether "inequality might be objectionable in itself, apart from whatever effects it might have."¹ Scanlon's answer is that distributive inequality is unjust when it is arbitrary in the sense that "the institutional mechanisms generating it cannot be justified in the right way."²And he proposes the following principle for determining when such a justification is available: "justifiable inequalities must either be unavoidable consequences of the exercise of important personal liberties or result from features of the economic system that are required in order for it to function in a way that benefits all."³

The reference to personal liberties in Scanlon's principle is intended to accommodate cases, most famously discussed by Robert Nozick, where inequalities may be justified because their elimination would wrongfully interfere with personal liberties.⁴ But while much has been written on this topic, I will set it aside here, and focus instead on the second part of Scanlon's principle, which applies even when personal liberties are not at stake. In such cases, Scanlon notes that his principle imposes a "weaker requirement" on the justification of inequality than John Rawls's difference principle, which "requires that a system that

¹ Scanlon 2018, p. 136.

² Ibid., p. 138.

³ Ibid., p. 151.

⁴ Nozick 1974

generates inequalities must not only benefit all, but must benefit those who have less as much as possible."⁵ And while he acknowledges that the "rationale" for his principle "leads naturally to something very much like Rawls's stronger principle," he maintains that "it is not necessary, in order to condemn existing levels of inequality, to follow the argument through to this conclusion."⁶

The rationale Scanlon has in mind is the familiar idea that inequalities may sometimes be justified by their efficiency or "economic productivity."⁷ Our institutions, after all, do not "merely allocate some independently existing set of benefits."⁸ They also facilitate their production. And though Scanlon insists that institutions that generate inequality stand in need of justification, he allows that not just considerations of personal liberty but also gains to efficiency in the production of benefits can play this justificatory role. Still, on Scanlon's view, not just any efficiency gain will do. Institutions that generate inequality are never justified merely by their production of a large quantity of benefits, but only when they are "productive in a way that benefits all."⁹ So if existing inequalities are in themselves unjust, it is because they are not efficient inequalities in this way: "they arise from features of our economic system that benefit only the rich."¹⁰

Scanlon's goal, then, is to come up with a principle that allows us to explain why existing inequalities are unjust without appealing to claims about their further effects, relying

⁵ Scanlon 2018, p. 142.

⁶ Ibid., p. 151.

⁷ Ibid., p. 140.

⁸ Ibid., p. 141.

⁹ Ibid., p. 143.

¹⁰ Ibid., p. 142.

on "anything as demanding as Rawls's Difference Principle," or denying that inequalities can sometimes be justified by their efficiency.¹¹ But while I share Scanlon's conviction that existing inequalities are unjust, I will argue that his principle cannot explain why. In particular, I will show that if one accepts, with Scanlon, both that institutions that generate inequality are sometimes for that very reason unjust, and that inequality is justified on grounds of efficiency in cases where it benefits everyone, then—given only a few very mild further assumptions—one must, on pain of inconsistency, also accept that sufficiently large gains to total benefits can justify inequality. One's objection to existing inequalities cannot be that the institutions that generate them fail to benefit everyone but must instead be that such institutions fail to produce enough total benefits to offset their unequal distribution.

My argument for this rather surprising conclusion proceeds in three steps. First, I show that given only the assumption that leveling down is unjust—an assumption Scanlon himself endorses—and a pair of weak and widely accepted further claims about the formal properties of justice, Scanlon's principle entails the same assessments of comparative justice as the difference principle in cases where no other demands of justice are relevant. Second, I build on this result to show that, given only one further weak and uncontroversial claim about equality, Scanlon's principle turns out to be inconsistent in the sense that it sometimes yields logically contradictory pairs of judgments in cases where one institution is more equal than another but the latter is better for the very worst off. Finally, I consider how we might modify Scanlon's principle to avoid this inconsistency while retaining a commitment to the core idea that efficiency can justify inequality at least when it benefits everyone, and show that the only consistent principle of this type allows sufficiently large gains in the production of total benefits to justify inequality. So unless one denies the injustice of leveling down or at

¹¹ Ibid., p. 138.

least one of three relatively uncontroversial properties of justice and equality, one cannot appeal to Scanlon's principle to explain why distributive inequalities are in themselves unjust. If institutions that generate inequality are justified when they are efficient in a way that benefits everyone, then they may also be justified by their production of a sufficient number of total benefits—regardless of how these benefits are distributed.

2. Scanlon's Principle and the Difference Principle

Scanlon proposes his principle as an alternative to Rawls's difference principle. But before we explore the relation between the two principles, we must first discard a superficial difference between them, namely, that Scanlon's allows "the exercise of important personal liberties" to justify inequality, but Rawls's does not.¹² This difference is superficial because Rawls himself allows that the exercise of individual liberties may justify inequality. On Rawls's view, the difference principle only applies once the demands of two other principles are met: one requiring the protection of basic liberties, and one ensuring fair equality of opportunity.¹³ It states that, in circumstances where all such prior demands of justice are satisfied, justice requires that "social and economic inequalities… are to be to the greatest benefit of the least advantaged members of society."¹⁴ Under these same circumstances, the reference to personal liberties drops out of Scanlon's principle, since we may assume that the exercise of liberties is already protected. Given this background assumption—which I will adopt throughout the rest of this paper—Scanlon's principle reads: "justifiable inequalities must… result from features of the economic system that are required in order for it to

¹⁴ Ibid.

¹² Ibid., p. 151.

¹³ Rawls 2001, pp. 42-43.

function in a way that benefits all."15

Scanlon is clear in his exposition of his principle that he is concerned not only with the injustice of deviations from perfect equality, but also with the injustice of there being greater rather than lesser inequality.¹⁶ He is also explicit that his principle, like the difference principle, relativizes evaluations of justice to the feasible set of institutions, holding that an institution that generates inequality fails to be efficient in a way that benefits everyone, and is therefore unjust, when there is some feasible institution that would be more equal without being worse for everyone.¹⁷ I therefore suggest that, at a first pass, we interpret Scanlon's principle as follows (keeping in mind that the principle, as stated, applies only in circumstances where all other demands of justice are satisfied, such that the justice of an institution depends entirely on the distribution of benefits it produces):

¹⁵ Scanlon 2018, p. 151. As Scanlon notes, one remaining difference is that Rawls is concerned only with certain "basic liberties," whereas Scanlon is concerned with "important personal liberties" more generally; ibid., fn. 34. This difference is unimportant for present purposes, since we may assume that all relevant liberties are protected no matter which are relevant.

¹⁶ Ibid., pp. 138, 144, 151.

¹⁷ Ibid, p. 142. A complete specification of Scanlon's principle would therefore require us to provide some account of when institutions are feasible. Scanlon provides us with little help here, stating only that his principle "makes conclusions about fairness depend on complex empirical questions about how economic institutions do function and how they could function if arranged differently"; ibid. Thankfully, we need not settle such issues here, since my argument goes through regardless of how one understands feasibility. Scanlon's Principle (first pass): An institution is unjust if there is some feasible institution that is more equal without being worse for everyone. Otherwise it is just.

The difference principle, on the other hand, says that (in these same circumstances):

Difference Principle: An institution is unjust if there is some feasible institution that is better for the very worst off. Otherwise it is just.

To see the contrast between these principles, let us refer to institutions by the distributions of benefits they produce, and represent these distributions as ordered *n*-tuples, such that <4, 5, 6> represents an institution that generates a distribution in which one individual has a benefit level of 4, a second a benefit level of 5, and a third a benefit level of 6.¹⁸ Suppose now that there are only three feasible institutions, <4, 4, 4>, <1, 5, 9>, and <5, 5, 6>, and that <1, 5, 9> is the least equal of the three. In this case, the Difference Principle claims that <4, 4, 4> and <1, 5, 9> are unjust, and that <5, 5, 6> is just, because <5, 5, 6> is better for the very worst off than the other two. Scanlon's Principle states that <1, 5, 9> is unjust because <4, 4, 4> is more equal without being worse for everyone, such that <1, 5, 9> generates an inequality that fails to benefit everyone. Unlike the Difference Principle, however, it implies that both <4, 4, 4> and <5, 5, 6> are just.

Now, one immediate objection to <4, 4, 4> is that it is worse for everyone than

¹⁸ Though Scanlon himself is concerned with inequalities in income and wealth, I here speak more generically of benefits in order to remain neutral on the vexed question of which sorts of benefits—income and wealth, well-being, capabilities, whatever—serves as the appropriate "currency of egalitarian justice"; Cohen 1989. I assume only that these benefits are numerically representable, and admit the relevant sorts of interpersonal comparisons. another feasible alternative: <5, 5, 6>. At various points, Scanlon recognizes this objection as legitimate. He says that "leveling down"—choosing a more equal distribution over one that is better for everyone—is "irrational,"¹⁹ "pointless,"²⁰ and "objectionable,"²¹ and that when inequality is justified "the pursuit of greater *equality* would be unjustifiable."²² Since it is uncontroversial to hold that leveling down is unjust,²³ I therefore suggest that we interpret Scanlon as claiming that institutions are unjust not only when they generate inequalities that fail to benefit everyone, but also when they generate distributions (equal or unequal) that are worse for everyone. Putting this together, we get:

Scanlon's Principle: An institution is unjust if there is some feasible institution that is either (i) better for everyone or (ii) more equal without being worse for everyone. Otherwise it is just.

So, given the feasible set consisting of <4, 4, 4>, <1, 5, 9>, and <5, 5, 6>, it follows by clause (i) that <4, 4, 4> is unjust (since <5, 5, 6> is better for everyone than the "leveled down" <4, 4, 4>), and it follows by clause (ii) that <1, 5, 9> is unjust (since <4, 4, 4> is more equal without being worse for everyone than <1, 5, 9>). In this case, then, Scanlon's Principle agrees with the Difference Principle. But there may be others where it does not. For example, if we take a feasible set consisting of <3, 6, 6> and <5, 5, 100>, then while the Difference Principle says that <3, 6, 6> is unjust and <5, 5, 100> is just, Scanlon's Principle

¹⁹ Scanlon 2018, p. 3.

²⁰ Ibid., p. 6.

²¹ Ibid., p. 28.

²² Ibid., p. 139.

²³ See, e.g., Parfit 1997.

will only claim that <3, 6, 6> is unjust if it is less equal than <5, 5, 100>. But if we instead endorse the intuitive verdict that <3, 6, 6> is more equal than <5, 5, 100>, then Scanlon's Principle disagrees with the Difference Principle in this case: it entails that <3, 6, 6> is just and that <5, 5, 100> unjust, since <3, 6, 6> is more equal without being worse for everyone.

So far, we have been considering categorical principles that make claims about what is just and unjust, rather than comparative principles that make claims about what is more just than what. The difference principle is, however, often described in comparative terms, as claiming that if one institution is better for the very worst off than a second, then the first is more just. To bridge the gap between these two interpretations, the following will do:

Universality: If institution x is more just than institution y in one feasible set, then x is more just than y in any logically possible feasible set.

Universality expresses the uncontroversial thought that, unlike categorical judgments of just or unjust, comparative judgments of more or less just are universal: they are not relativized to a feasible set. This is a basic principle of our moral and political reasoning. If one institution is more just than another, then whether or not some third institution is feasible should not affect this comparison—for example, if welfare state capitalism is more just than laissez-faire capitalism, then this should be true regardless of whether communism is feasible. Now, to be fair, there are some who object to this principle (or structurally similar principles that apply in other normative domains) on the grounds that, in some cases, the feasibility of options may affect an institution's justice-relevant properties. For example, which options are feasible may affect the size of individuals' complaints under those institutions, since the size of a complaint may depend on how badly off an individual is relative to feasible alternatives.²⁴ But since all the principles we are concerned with here pick out benefit levels as the only justice-relevant properties, and since benefit levels may be specified independently of the feasibility of other institutions, this sort of objection to Universality does not apply to them. If the fact that institution x is better for everyone than y, more equal without being worse for everyone than y, or better for the very worst off than y, makes x more just than y in one feasible set, then the same should be true in any feasible set. And, in the present context, this is all Universality says.

Once we recognize Universality, we may transform categorical principles into comparative ones through the following trick. Consider a feasible set of two institutions, xand y, such that x is either better for everyone than y, or more equal than y without being worse for everyone. In each case, Scanlon's Principle implies that x is just and that y is unjust, and, therefore, that x is more just than y. But now, by Universality, it follows that if xis more just than y in this feasible set, it must be more just than y in any logically possible feasible set. So Scanlon's Principle and Universality together imply:

Scanlon's Comparative Principle: In every logically possible feasible set, institution x is more just than institution y if (i) x is better for everyone than y, or (ii) x is more equal than y without being worse for everyone.

For example, in a feasible set containing only <4, 4, 4>, <3, 6, 6>, and <5, 5, 100>, clause (i) of Scanlon's Comparative Principle implies that <5, 5, 100> is more just than <4, 4, 4> (because it is better for everyone), and clause (ii) implies that <4, 4, 4> is more just than <3, 6, 6> (because it is more equal without being worse for everyone). And if we appeal to

²⁴ Tungodden and Vallentyne 2005, p. 147. Compare Sen's 1997, pp. 752-755 discussion of "menu-dependence."

Universality and apply the same trick to the Difference Principle to derive:

Comparative Difference: In every logically possible feasible set, institution x is more just than institution y if x is better for the very worst off than y.

then, in this feasible set, Comparative Difference will agree with these two judgments, and claim further that <5, 5, 100> is more just than <3, 6, 6>.

We now introduce a second formal property of justice:

Transitivity: If institution x is more just than institution y, and institution y is more just than institution z, then x is more just than z.

Like Universality, Transitivity is a basic principle of our moral and political reasoning.²⁵ If one institution is more just than another, and this institution is more just than a third, then the first institution is more just than the third—for example, if welfare state capitalism is more just than laissez-faire capitalism, and laissez-faire capitalism is more just than communism, then welfare state capitalism is more just than communism. And if we add Transitivity to our previous claims, then Scanlon's Comparative Principle will agree with Comparative Difference about the ranking of options in the feasible set <5, 5, 100>, <4, 4, 4>, and <3, 6, 6>. For, as we have just seen, Scanlon's Comparative Principle entails that <5, 5, 100> is more just than <4, 4, 4>, and that <4, 4, 4> is more just than <3, 6, 6>. So, by Transitivity, it follows that <5, 5, 100> is more just than <3, 6, 6>, too.

In fact, we now have all the conditions we need to show that Scanlon's Comparative Principle *always* agrees with Comparative Difference. For if <5, 5, 100> is more just than <3, 6, 6> when <4, 4, 4> is feasible, then it follows from Universality that <5, 5, 100> is always

²⁵ Though see Temkin 2011.

more just than <3, 6, 6>. And, more generally, whenever institution x is better for the worst off than institution γ , it will follow from Scanlon's Comparative Principle that x is more just than y. For if x is better for everyone than y, it will follow by clause (i) that x is more just. And if x is not better for everyone than y, then we will always be able to add a perfectly equal institution e to the feasible set whose members are worse off than the worst off individual under x but better off than the worst off individuals under y, and run the same argument as in the case where x is $\langle 5, 5, 100 \rangle$, e is $\langle 4, 4, 4 \rangle$, and y is $\langle 3, 6, 6 \rangle$: x is better for everyone than e and so is more just than e by clause (i), e is more equal than y without being worse for everyone than y and so is more just than y by clause (ii), so by Transitivity x is more just than y, and by Universality this holds not only in this feasible set, but in all logically possible feasible sets. The case where x is $\langle 5, 5, 100 \rangle$, y is $\langle 3, 6, 6 \rangle$, and e is $\langle 4, 4, 4 \rangle$ provides just one example of this general line of reasoning. To take another: if x is <2, 2, 8>, and y is <1, 3, 5>, then we can imagine a logically possible feasible set that includes these two options along with <1.5, 1.5, 1.5>. It follows that <2, 2, 8> is more just than <1.5, 1.5, 1.5> by clause (i), that <1.5, 1.5, 1.5 is more just than <1, 3, 5> by clause (ii), and therefore that <2, 2, 8> is more just than <1, 3, 5> by Transitivity. And, by Universality, this holds regardless of whether or not <1.5, 1.5, 1.5 is actually feasible.

I borrow the insight of this informal proof—that whenever institution x is better for the worst off than institution y, we can pick out a perfectly equal institution that, according to Scanlon's Comparative Principle, is less just than x but more just than y—from Bertil Tungodden and Peter Vallentyne, who provide a formal proof of a very similar result.²⁶ Although it does not get us all the way to the claim that Scanlon's Principle implies the (noncomparative) Difference Principle, it does show that, given Universality and Transitivity,

²⁶ Tungodden and Vallentyne 2005, Result 3.

Scanlon's Principle implies Comparative Difference.²⁷ And if we add to this the claim that a society is unjust when it is less just than some feasible alternative, then we do arrive at the Difference Principle itself. This further claim, however, is not needed in what follows. Nor, strictly speaking, do we need Transitivity. For consider:

Acyclicity: If institution x is more just than institution y, and institution y is more just than institution z, then z is not more just than x.

According to Acyclicity, if one institution is more just than a second, and this second institution is more just than a third, then, even if the first institution is not more just than the third (as Transitivity implies) the third is at least not more just than the first—if welfare state capitalism is more just than laissez-faire capitalism, and laissez-faire capitalism. Acyclicity than communism, then communism is not more just than welfare state capitalism. Acyclicity is therefore weaker than Transitivity, and if we replace the latter with the former, then precisely the same informal proof can be given to show that whenever one institution is better for the very worst off than a second, Scanlon's Principle implies that the second institution is not more just than *y*, so *x* is more just than *y*," we will now say "*x* is more just than *e*, *e* is more just than *y*, so *y* is not more just than *x*." This, as we will now see, is enough to trap Scanlon's Principle in inconsistency, at least if we accept one very weak claim

²⁷ Technically, this implication has relied on two further assumptions: that a just institution is more just than an unjust one, and that a perfectly equal institution (under which everyone has the same) is more just than one producing some inequality (under which some have more than others). I take both of these assumptions to be analytic, but they are worth flagging here to ensure that the argument is airtight. about equality. And since Acyclicity is even less controversial than Transitivity—though even it has its detractors²⁸—it is all we need rely on in what follows.

3. The Inconsistency of Scanlon's Principle

In the last section, I showed that Scanlon's Principle, conjoined with Universality, implies Scanlon's Comparative Principle, the second clause of which I will now isolate and call:

Equality Priority: In every logically possible feasible set, if institution x is more equal than institution y without being worse for everyone, then x is more just than y.

I then showed that if we add Transitivity, Scanlon's Principle furthermore implies Comparative Difference, and that if we weaken Transitivity to Acylicity it still implies a principle according to which our evaluations of comparative justice never disagree with Comparative Difference, which I will now label:

Worst Off Priority: In every logically possible feasible set, if institution x is better for the very worst off than institution y, then y is not more just than x.

Though these two principles may seem like kindred spirits, they unfortunately imply inconsistent verdicts in an important range of cases. To see this, recall the earlier feasible set we used to illustrate the disagreement between Scanlon's Principle and the Difference Principle: one consisting of <5, 5, 100> and <3, 6, 6>. The key feature of this example was that, though the former is better for the worst off, the latter is intuitively more equal without being worse for everyone, such that—if we grant this intuitive claim—the Difference Principle deems <5, 5, 100> just and <3, 6, 6> unjust while Scanlon's Principle entails the

²⁸ See, again, Temkin 2011.

reverse. At the time, this divergence did not seem problematic for Scanlon's Principle, which is, after all, intended as an alternative to the Difference Principle. But now that we have seen that Scanlon's Principle implies not only categorical claims about justice but also the comparative claims I have labeled Equality Priority and Worst Off Priority, a problem emerges. For if <3, 6, 6> is indeed more equal than <5, 5, 100> then it follows by Equality Priority that <3, 6, 6> is more just than <5, 5, 100> (since the former is more equal without being worse for everyone), while it follows from Worse Off Priority that <3, 6, 6> is not more just than <5, 5, 100> (since the latter is better for the very worst off). In other words, since Scanlon's Principle (conjoined with Universality and Acyclicity) implies both Equality Priority and Worst Off Priority, it implies further that if <3, 6, 6> is more equal than <5, 5, 100>, then <3, 6, 6> both *is* and is *not* more just than <5, 5, 100>. And this is inconsistent.

We can avoid this particular inconsistency, of course, simply by denying that <3, 6, 6> is more equal than <5, 5, 100>. But the example does not matter. Scanlon's Principle will imply inconsistent verdicts in any case where x is more equal than y without being worse for everyone, but y is better for the very worst off, since in all such cases Equality Priority will imply that x is more just than y while Worse Off Priority will imply that x is not more just than y. Assuming we accept Universality and Acyclicity, then, the only way to rescue Scanlon's Principle from inconsistency is to deny that such cases ever occur, by endorsing:

Equality-Worst Off Link: There is no logically possible case where both (i) institution x is more equal than institution y without being worse for everyone than y, and (ii) y is better for the very worst off than x.

Equality-Worst Off Link implies, for example, that <3, 6, 6> cannot be more equal than <5, 5, 100>. And more generally, it is the weakest claim we can appeal to in order to ensure that

there are no cases where our judgments about equality generate the inconsistency between Equality Priority and Worst Off Priority just mentioned.

Tungodden and Vallentyne notice this inconsistency, but they do not investigate the plausibility of Equality-Worst Off Link.²⁹ In previous work, however, I have provided three arguments against Equality-Worst Off Link myself.³⁰ The first argument appeals to the intuitive notion of a deviation from perfect equality. Even if one finds the comparison between <3, 6, 6> and <5, 5, 100> unconvincing, it is easy to construct more extreme examples. For instance, it certainly seems that <3, 6, 6, 6> deviates less from perfect equality, and is therefore more equal, than <5, 5, 100, 999>---or, to take one of the cases I employ, that <1.99, 2.01, 2.01, 2.01, 2.01> deviates less from perfect equality and is therefore more equal than <2, 2, 100, 1000, 10000>—since the former is almost perfectly equal, and the latter is far from it. And this is enough to disprove Equality-Worst Off Link. My second argument appeals to formal properties that are widely endorsed by economists ²⁹ Tungodden and Vallentyne 2005, p. 149. Though Tungodden and Vallentyne do not have Scanlon in mind, they do show, in effect, that if we strengthen clause (i) of Scanlon's Comparative Principle to the claim that x is more just than y when it is better for some and worse for none (and not merely when it is better for all) than y, and weaken clause (ii) of Scanlon's Comparative Principle ("Equality Priority") accordingly, then, given just a few further claims, the principle runs into a similar inconsistency; Ibid., Result 2. The advantage

of the argument given here is that it does not depend on this strengthening of clause (i), which Scanlon does not accept.

³⁰ Barrett 2018. In that paper I actually argue against a weaker claim than Equality-Worst Off Link, whose rejection implies the rejection of Equality-Worst Off Link. This will be important later. concerned with measuring distributive inequality. The key property here is Pigou-Dalton (also known as the "transfer axiom"), which holds that if we transfer some fixed quantity of benefits from a better off to a worse off person, without losing or gaining any benefits or affecting anyone else, this improves equality. I show that Pigou-Dalton implies the falsity of Equality-Worst Off Link given only a couple other weak and widely accepted assumptions a result that is unsurprising given that all standard metrics for measuring inequality, the most famous of which is the Gini index, contradict Equality-Worst Off Link. My third argument appeals to the claim that the equality of an entire distribution does not depend entirely on whether its two worst off members are equally well off. This, I prove, entails the failure of Equality Worst-Off link, again, given only weak and widely accepted further assumptions.

Even without going into the details of these latter two proofs, it is clear that Equality Worst-Off Link is a highly implausible claim. As Dennis McKerlie puts the worry: "If we care about equality it is plausible to think that we object to inequality between any two groups. How can we get from this starting-point to the conclusion that we should assess inequality only in terms of its effect on the worst off?"³¹ And the obvious answer is: we can't. Regardless of what we think about their relative justice, it is ridiculous to deny that <1.99, 2.01, 2.01, 2.01, 2.01> is more equal than <2, 2, 100, 1000, 10000> just because the latter is marginally better for the very worst off, and any criterion of equality that claims otherwise is no criterion *of equality* at all. So we should reject Equality Worst-Off Link, and with it, Scanlon's Principle. After all, we have seen that unless one denies that leveling down is unjust (something that Scanlon accepts, and which I have therefore included in my statement of Scanlon's Principle), or denies Universality or Acyclicity (which are both plausible and widely accepted formal properties of justice), it follows from the rejection of

³¹ McKerlie 1994, p. 33.

Equality-Worst Off Link that Scanlon's Principle yields inconsistent verdicts. And any principle that yields inconsistent verdicts must be rejected.

4. Revising Scanlon's Principle

The failure of Scanlon's Principle is interesting not only in its own right, but also because it raises the broader question of whether there is some similarly motivated principle that can explain when inequalities are in themselves unjust without collapsing into inconsistency.³² In particular, we have seen that Scanlon proposes his principle as a less demanding alternative to the Difference Principle that accommodates the core idea that while institutions that generate inequality stand in need of justification, gains to efficiency in the production of benefits may play this justificatory role. Scanlon runs into trouble because of the way he fleshes this idea out. He claims that inequalities are unjust unless they meet a very stringent efficiency requirement: they must benefit everyone. But there are various ways we might revise Scanlon's Principle while retaining the core idea that only efficient inequalities are justified by relaxing this requirement and allowing that efficiency can justify inequality not only when it benefits everyone, but in a wider range of cases. And it is worth investigating whether there is some principle of this sort that can provide a consistent explanation of whether there is not principle of this sort that can provide a consistent explanation of

In this section, I undertake this investigation. I show that while it is indeed possible to rescue Scanlon's Principle from inconsistency in this way, one can do so only by allowing sufficiently large gains in the production of total benefits to justify inequality, regardless of how these benefits are distributed. In other words, if one holds, with Scanlon, that inequality is justified when it is efficient in a way that benefits everyone, then one must also hold that

³² Thanks to an anonymous reviewer for urging me to say more about this broader question.

sufficiently large gains to Utilitarian efficiency justify inequality, where:

Utilitarian efficiency: institution x is more Utilitarian efficient than institution y if x produces a greater total sum of benefits than y.³³

And one must therefore claim that inequalities are unjust not when the institutions that generate them fail to benefit everyone, but when such institutions fail to produce enough total benefits to make up for their unequal distribution of these benefits.

I defend this conclusion by first surveying what other conceptions of efficiency we might draw on to revise Scanlon's Principle, and then showing that any revision that doesn't appeal to Utilitarian efficiency yields inconsistent verdicts in many of the same cases as Scanlon's Principle itself. Now, in the literature, the conception of efficiency that Scanlon's Principle invokes is known as *Weak Pareto* efficiency, where:

Weak Pareto: institution x is more Weak Pareto efficient than institution y if x is better for everyone than y. ³⁴

This is often contrasted with Strong Pareto efficiency, where:

³³ Although Utilitarians are concerned with producing a greater total sum of *well-being*, I here generalize this idea to a concern with producing a greater total sum of whichever sorts of benefit serves as the relevant "currency" of justice; see fn. 18 above. I also set aside the familiar question of whether we should be concerned with producing greater total benefits or greater average benefits. For our purposes, it is enough to note that, on either view, gains to efficiency always involve the production of greater total benefits in cases where we are concerned with a fixed population, and that these are the only cases we need consider here. ³⁴ See, e.g., Sen 2018.

Strong Pareto: institution x is more Strong Pareto efficient than institution y if x is better for some and worse for none than y. 35

Strong Pareto is stronger than Weak Pareto in the sense that a gain to Weak Pareto efficiency is always a gain to Strong Pareto efficiency, but the reverse does not hold. For example, <5, 8> is both more Weak Pareto and more Strong Pareto efficient than <4, 5> because it is better for everyone, and, therefore better for some and worse for none. But <5, 8> is more Strong Pareto efficient than <5, 6> without being more Weak Pareto efficient, since it is better for some and worse for none without being better for everyone.

One interesting feature of Strong Pareto is that it makes our judgments of efficiency depend not only on the overall shape of the distribution of benefits, but also on the particular individuals who occupy the positions in that distribution. For example, whereas <5, 8> is more Strong Pareto efficient than <5, 6> (since it is better for some and worse for none), it is not more Strong Pareto efficient than <6, 5> (since it is worse for the first individual), even though <5, 6> and <6, 5> are *permutations* of each other in the sense that each produces the same distribution of benefits assigned to different individuals. If we wish to appeal to a stronger conception of efficiency than Strong Pareto, we may therefore deny that such permutations ever affect efficiency. This brings us to:

Suppes: institution x is more Suppes efficient than institution y if x is better for some and worse for none than y or than some permutation of y.³⁶

Suppes agrees with Strong Pareto that whenever an institution is better for some and worse

³⁵ Ibid.

³⁶ I adapt this condition from Suppes 1966.

for none, it is more efficient. But it goes further and claims that <5, 8> is more Suppes efficient than <6, 5> because, even though it is not better for some and worse for none, it is better for some and worse for none than some permutation of <6, 5>, namely, <5, 6>.

Although Suppes efficiency denies that permutations ever make a difference to efficiency, it still deems it relevant how benefits are distributed more generally. For example, while <8, 5> is more Suppes efficient than <7, 5> and than <5, 7>, it is not more Suppes efficient than <6, 6>, even though <7, 5>, <5, 7> and <6, 6> are *redistributions* of each other in the sense that each contains the same total sum of benefits distributed differently. If we wish to appeal to an even stronger conception of efficiency than Suppes, we may therefore claim not only that permutations never make a difference to efficiency, but also that redistributions never make this difference. This yields:

Kaldor-Hicks: institution x is more Kaldor-Hicks efficient than institution y if x is better for some and worse for none than y or than some redistribution of y.³⁷

So, for example, while <5, 8> is not more Weak Pareto, Strong Pareto, or Suppes efficient than <6, 6>, <5, 8> is more Kaldor-Hicks efficient than <6, 6> because it is better for some and worse for none than some redistribution of <6, 6>, namely, <5, 7>.

A moment's reflection reveals, however, that Kaldor-Hicks efficiency is equivalent to Utilitarian efficiency. Since redistributions, by definition, never increase or decrease the total 37 Kaldor 1939, Hicks 1939. Kaldor-Hicks efficiency is typically articulated as the claim that x is more efficient than y if it would be logically possible to redistribute the *material goods* in x in such a way that this would make some better off and none worse off with respect to *well-being*. But, as before, I generalize this to a concern whatever sort of benefit is relevant to justice; see fns. 18, 33 above.

sum of benefits, it is logically possible to redistribute benefits in such a way that x is better for some and worse for none than y if and only if there are more total benefits in x than in y.³⁸ So there are only four distinct conceptions of efficiency in the literature, and—since there does not appear to be conceptual room for a fifth—I will assume that we have only four conceptions to work with: Weak Pareto, Strong Pareto, Suppes, and Utilitarian. With these conceptions in hand, we may relax the requirement set by Scanlon's Principle in a number of ways. For example, whereas Scanlon's Principle holds that an institution is unjust unless it is more Weak Pareto efficient (better for everyone) than any more equal feasible institution, we may now claim that gains to Strong Pareto, Suppes, or Utilitarian efficiency also justify inequality. Or we may adopt a variant of any of these principles on which inequality is justified not by any gain to the relevant sort of efficiency, but only by gains that are sufficiently large either in absolute terms or in relation to the degree of inequality at issue. And then there are hybrid principles, on which, say, any gain to Strong Pareto efficiency justifies inequality, but only sufficiently large gains to Utilitarian efficiency

With this survey complete, I am now ready to present my argument that any viable modification to Scanlon's Principle must fall in the class of principles that allows sufficiently large gains to Utilitarian efficiency to justify inequality—a class that includes as a limit case a principle, once defended by Henry Sidgwick, on which *any* marginal gain to Utilitarian efficiency is sufficiently large to justify inequality.³⁹ To begin, note that any principle that doesn't fall in this class can only permit gains to Weak Pareto, Strong Pareto, and Suppes efficiency to justify inequality. And, since gains to Weak and Strong Pareto efficiency are ³⁸ Technically, this equivalency only holds when comparing two fixed population sizes. But

these are the only cases I am concerned with here; see fn. 33 above.

³⁹ Sidgwick 1981, pp. 416-417.

always gains to Suppes efficiency, it follows that all such principles are *Suppes-maximal*, in the sense that they at most allow gains to Suppes efficiency to justify inequality. Suppes-maximal principles include a range of principles anywhere from Scanlon's Principle on which gains to Suppes efficiency only justify inequality when they are also gains to Weak Pareto efficiency, to a principle on which any gain to Suppes efficiency justifies inequality. But they all agree, first, that gains to Weak Pareto efficiency justify inequality, and, second, that gains to Utilitarian efficiency never justify inequality unless they are also gains to Suppes efficiency.

The task of demonstrating that the only revisions to Scanlon's Principle that allow it to avoid inconsistency must permit sufficiently large gains to Utilitarian efficiency to justify inequality therefore reduces to showing both that all Suppes-maximal principles run into inconsistency, and that some principles that appeal to Utilitarian efficiency do not. And sure enough, this is what we find. If we take the original version of Scanlon's Principle:

Scanlon's Principle: An institution is unjust if there is some feasible institution that is either (i) better for everyone or (ii) more equal without being worse for everyone. Otherwise it is just.

and we continue to endorse the injustice of leveling down articulated by clause (i), but relax the efficiency requirement set by clause (ii) in any way that continues to imply that inequalities are unjust when there is some feasible institution that is more equal without being less Suppes efficient, then given only Universality, Acyclicity, and the rejection of a revised version of Equality-Worst Off Link, we are committed to inconsistent verdicts in many of the same cases as before. But if we revise clause (ii) in a way that permits sufficiently large gains to Utilitarian efficiency to justify inequality, then this inconsistency need not arise.

It would be possible for me to prove this by constructing a parallel argument to the

one of the last two sections. I would first show how we can appeal to Universality to convert any Suppes-maximal principle into a comparative principle with the same first clause as Scanlon's Comparative Principle but whose second clause instead implies:

Revised Equality Priority: if institution x is more equal than institution y without being less Suppes Efficient than y, then x is more just than y.

I would then show how if we add Transitivity or Acylicity, we can employ precisely the same informal proof as before—involving picking out a perfectly equal institution *e*—to derive Comparative Difference or the weaker version of it I have called Worst Off Priority. And I would finally point out that Revised Equality Priority and Worst Off Priority yield inconsistent verdicts in cases where *x* is more equal than *y* without being less Suppes efficient than *y* even though *y* is better for the very worst off. But it will perhaps be more illuminating instead to illustrate how to derive this inconsistency with an example, so that we can see both why it arises and why appealing to Utilitarian efficiency lets us avoid it.

To generate such an example, we need only find a case that falsifies:

Revised Equality-Worst Off Link: There is no logically possible case where (i) institution x is more equal than institution y without being less Suppes efficient than y, and (ii) y is better for the very worst off than x.

Technically, our earlier rejection of Equality-Worst Off Link does not imply the rejection of Revised Equality-Worst Off Link, since that condition's clause (i) only referred to cases where an institution x is more equal than y without being less Weak Pareto efficient (worse for everyone) than y. But the examples I chose to illustrate the intuitive case against Equality-Worst Off Link—those where x is <3, 6, 6, 6> and y is <5, 5, 100, 999>, or x is <1.99, 2.01,

2.01, 2.01, 2.01> and *y* is <2, 2, 100, 1000, 10000>—also falsify Revised Equality-Worst Off Link, since, in each case, *x* is intuitively more equal than *y* without being less Suppes efficient than *y*. And the two proofs that I mentioned earlier entail not only the rejection of Equality-Worst Off Link but also the rejection of Revised Equality-Worst Off Link.⁴⁰ So there is no reason to reject the former without also rejecting the latter. And, as we will now see, this is enough to generate an inconsistency for any Suppes-maximal revision to Scanlon's Principle.

Although any pair of cases that falsifies Revised Equality-Worst Off Link will do, let us assume, for concreteness, that <3, 6, 6> is indeed more equal than <5, 5, 100> even though it is not more Suppes efficient. This pair of institutions should be familiar: they make up the feasible set we used first to demonstrate the conflict between Scanlon's Principle and the Difference Principle, and then to draw out the inconsistency of Scanlon's Principle. We now show that no Suppes-maximal revision of Scanlon's Principle can save us from the inconsistent pair of verdicts that <3, 6, 6> both *is* and is *not* more just than <5, 5, 100>. The first step is to note that, because <5, 5, 100> is not more Suppes efficient than <3, 6, 6> (it is not better for some and worse for none than <3, 6, 6> or than any permutation of <3, 6, 6>), but the latter is *ex hypothesi* more equal, it follows from clause (ii) of Scanlon's principle or any Suppes-maximal revision that in a feasible set containing only these two institutions, <3, 6, 6> is just and <5, 5, 100> is unjust, such that the first is more just than the second. By Universality, we obtain the first verdict in our inconsistent pair: <3, 6, 6> is more just than <5, 5, 100> in any logically possible feasible set.

To derive the second verdict, we employ our earlier move of picking out a perfectly equal institution e, say, <4, 4, 4> that is better for the worst off member in <3, 6, 6> but worse for the worst off member of <5, 5, 100>. We then note that, in a feasible set

⁴⁰ Barrett 2018; see fn. 30 above.

containing only <5, 5, 100> and <4, 4, 4>, the former is just and the latter is unjust given clause (i) of Scanlon's Principle or any Suppes-maximal revision (since <4, 4, 4> is worse for everyone). And in a feasible set containing only <4, 4, 4> and <3, 6, 6>, the former is just and the latter is unjust given clause (ii) of Scanlon's Principle or any Suppes-maximal revision (since <3, 6, 6> is less equal than <4, 4, 4> without being more Suppes efficient, that is, without being better for some and worse for none than <4, 4, 4> or than any permutation of <4, 4, 4> is more just than <3, 6, 6>, such that, by Universality <5, 5, 100> is more just than <4, 4, 4> is more just than <3, 6, 6> in any logically possible feasible set, and <4, 4, 4> is more just than <3, 6, 6> in any logically possible feasible set. By Acyclicity, we therefore reach our second verdict: <3, 6, 6> is *not* more just than <5, 5, 100> in any logically possible feasible set. But this conflicts with our prior verdict that <3, 6, 6> *is* more just than <5, 5, 100>. So we have an inconsistency in this case, just as with the original version of Scanlon's Principle itself.

This technique for generating inconsistencies did not depend at any point on the example I chose, and so can be generalized to any case where Revised Equality-Worst Off Link fails such that (i) institution x is more equal than y without being less Suppes efficient than y, and (ii) y is better for the very worst off than x. Thus, we come to the advertised conclusion that any viable modification of Scanlon's Principle must allow not only gains to Suppes efficiency (which include gains to Weak and Strong Pareto) to justify inequality, but must also allow sufficiently large gains to Utilitarian efficiency to justify inequality. In the example at hand, for instance, inconsistency can be avoided only by rejecting either the step where we held that <3, 6, 6> is unjust in a feasible set consisting of it and <4, 4, 4>.

And since, in both cases, the former is less equal than the latter without being more Suppes efficient, the only efficiency-based explanation of this must appeal to Utilitarian efficiency namely, to the claim either that the greater inequality of <5, 5, 100> is justified by its production of so many more benefits than <3, 6, 6>, or that the greater inequality of <3, 6, 6> is justified by its production of so many more benefits than <4, 4, 4>.

To be clear, this does not push us all the way to Sidgwick's view that any gain to Utilitarian efficiency can justify inequality. Nothing I have said implies that one cannot instead hold that inequalities are only justified by gains to total benefits when these gains are sufficiently large in absolute terms or relative to the degree of inequality they engender. But I have shown that, if we think inequalities can be justified on grounds of efficiency when they benefit everyone, then we must at least allow sufficiently large gains to Utilitarian efficiency also to justify inequality, since holding that only Weak Pareto, Strong Pareto, or even Suppes efficiency can justify inequality leads us to inconsistent verdicts given a commitment only to the injustice of leveling down, to Universality and Acyclicity, and to the rejection of Revised Equality Worst-Off Link. We must allow, in other words, that inequality can sometimes be justified on the grounds that the institutions that generate it produce enough total benefits to offset this inequality, regardless of how these benefits are distributed.

5. Conclusion

There are many reasons one might object to existing inequalities in the distribution of social and economic benefits. As Scanlon and various others have argued, such inequalities may be unjust due to their effects (say, because they corrupt the political process or introduce objectionable status inequalities), or because they are produced in a way that violates some independent requirement of justice (say, that we do not discriminate on the basis of gender or race).⁴¹ And institutions that generate inequality might also be in themselves unjust, because such inequalities stand in need of further justification. One principle for explaining when such inequalities meet this standard is Rawls's Difference Principle: it claims that inequalities are unjust unless they benefit the worst off as much as possible. But this is an extremely demanding principle. So, on first glance, it is perhaps more plausible to endorse Scanlon's Principle, on which inequalities can be justified on grounds of their efficiency at least when they are efficient in a way that benefits everyone.

As we have seen, there is a rather close relation between these two principles. Given only a few further assumptions, Scanlon's Principle yields all the same comparative verdicts as the Difference Principle. But this convergence ultimately leads to the downfall of Scanlon's Principle, since it results in it collapsing under the weight of inconsistency in a range of cases where more equal institutions are worse for the very worst off. And attempts to tweak Scanlon's Principle by allowing efficiency to justify inequality in a wider range of cases without going all the way to the view that sufficiently large gains to Utilitarian efficiency-to total benefits-can justify inequality prove to be of no avail. Efficiency-based justifications of equality, it turns out, are subject to a certain sort of creep. Once we allow that efficient inequalities are justified in cases where they benefit everyone, we cannot stop there, but must admit that institutions that produce inequality are also justified when they produce enough total benefits. So if one agrees with Scanlon both that existing inequality is in itself unjust and that this inequality would be justified if it were efficient in a way that benefited everyone, then one's objection to existing inequality cannot be that it fails to benefit the worst off as much as possible (as Rawls claims), that it fails to benefit everyone (as Scanlon claims), or even, say, that it fails to benefit some without harming anyone. It

27

⁴¹ Scanlon 2018.

must instead be that while institutions that generate inequality can always in principle be justified by their production of a sufficient quantity of benefits, our current institutions do not create enough total benefits to make up for the inequality they produce.

REFERENCES

- Barrett, Jacob. 2018. Is Maximin Egalitarian? Synthese, Online First, 1-21.
- Cohen, G. A. 1989. On the Currency of Egalitarian Justice. Ethics, 99, 906-944.
- Hicks, J R. 1939. The Foundations of Welfare Economics. *The Economic Journal*, 49, 696-712.
- Kaldor, Nicholas. 1939. Welfare Propositions of Economics and Interpersonal Comparisons of Utility. *The Economic Journal*, 49, 549-552.
- McKerlie Dennis. 1994. Equality and Priority. Utilitas, 6, 25-42.
- Nozick, Robert. 1974. Anarchy, State, and Utopia. New York: Basic Books.
- Parfit, Derek. 1997. Equality and Priority. Ratio, 10, 202-221.
- Rawls John. 2001. Justice as Fairness: A Restatement. Cambridge: Belknap Press.
- Scanlon, T. M. 2018. Why Does Inequality Matter? New York: Oxford University Press.
- Sen, Amartya. 2018. Collective Choice and Social Welfare: An Expanded Edition. Cambridge:

Harvard University Press.

Sen, Amartya. 1997. Maximization and the Act of Choice. Econometrica, 65, 745-779.

Sidgwick, Henry 1981. The Methods of Ethics, 7th ed. Indianapolis: Hackett.

Suppes Patrick. 1966. Some Formal Models of Grading Principles. Synthese, 16, 284-306.

Temkin, Larry. 2011. Rethinking the Good. New York: Oxford University Press.

Tungodden, Bertil and Peter Vallentyne. 2005. On The Possibility of Paretian Egalitarianism. Journal of Philosophy, 102, 126-154.