

Sen on Sufficiency, Priority, and Equality

(April 6, 2013)

For *Contemporary Philosophy in Focus: Amartya Sen* (Cambridge University Press, 2010)

Peter Vallentyne, University of Missouri-Columbia

Abstract: I present a critical survey of Sen's work, and related work by others, on certain distribution-sensitive principles of justice. More specifically, I discuss three kinds of such principles: (1) sufficientarian principles, which require promoting the adequacy of individual benefits (non-poverty), (2) prioritarian principles, which require the promotion of individual benefits but with some kind of priority for the worse off, and (3) egalitarian principles, which require the promotion of equality of benefits. The focus is on presenting the issues and results in an intuitively accessible manner that highlights its importance for existing philosophical debates on the topic. Critical comments are only suggestive rather than fully defended.

1. Introduction

Amartya Sen's contributions to moral philosophy have been enormous. He has made seminal contributions to the measurement of freedom, the measurement of equality, the measurement of poverty, the debate over the kind of equality that is relevant to justice, and the debate over the respective roles of liberty, efficiency, and equality in a theory of justice. I shall here focus on Sen's work on distribution-sensitive principles of justice, and more specifically on his work on (1) sufficientarian principles, which require promoting the adequacy of benefits (non-poverty), (2) prioritarian principles, which require the promotion of individual benefits but with some kind of priority for the worse off, and (3) egalitarian principles, which require the promotion of equality of benefits.

Sen's work in this area tends to focus on the development of a framework of investigation rather than an extended defense of any particular principles. My goal here will therefore be to present Sen's work, and related work by others, in an intuitively accessible manner that highlights its importance for existing philosophical debates on the topic. The presentation will mainly be a survey, with critical comments being only suggestive rather than fully defended.

2. Background

Justice, as we shall understand it, is concerned with what benefits (net of specified burdens) we owe others. We shall leave open the nature of the relevant benefits (which might be income, wealth, well-being, or anything else). Sen, of course, has argued at length that the equalisandum for justice is capabilities (effective opportunities to function), but here we shall abstract from his commitment to capabilities, since that topic will be addressed in another chapter. Below, references to a person being better (or worse, or equally well) off than another should be understood as being assessed with respect to the relevant benefits (and not necessarily in terms of welfare). For simplicity, we shall assume that the value of given bundle of benefits is fully cardinally measurable and fully interpersonally comparable. This is indeed a simplifying assumption, but as Sen has emphasized, even if the bundles are not *fully* measurable in this way, they may be at least *partially* so measurable, and this may still generate significant implications.¹

Sen rightly insists that justice is pluralistic in that there are several distinct considerations that determine what is just.² He endorses at least some kind of impartiality condition, some kind of efficiency consideration, and some kind of distributive consideration. He also would endorse some kind of negative rights protecting against certain kinds of interference in one's life, but we shall ignore this important consideration in what follows, since it is not at the core of his work on

distribution-sensitive principles of justice.

In what follows, we shall focus on the ranking relation of being-at-least-as-just. A distribution of benefits is more just than another if and only if it is at least as just as the other, but not vice-versa. Two distributions are equally just if and only if each is at least as just as the other is. Except as noted, we shall make the following assumptions, which Sen and other social choice theorists typically make:

Ordering: The relation of being-at-least-as-just is an ordering: (1) It is *reflexive*: each distribution is at least as just as itself. (2) It is *transitive*: For any three distributions, x, y, and, z, if x is at least as just as y, and y is at least as just as z, then x is at least as just as z. (3) It is *complete*: for any two distributions, at least one of them is at least as just as the other.

Reflexivity is entirely uncontroversial, transitivity is somewhat controversial³, and completeness is moderately controversial. Nonetheless, for simplicity, we shall here assume them throughout.

A second condition that we shall generally assume is:

Anonymity: If the pattern of benefits in one distribution is a permutation of the pattern of benefits in another distribution (i.e., the same pattern of benefits but with the benefits reassigned to different people), then the two distributions are equally just.

Anonymity is an impartiality condition. It holds that (assuming, as we shall, that there is a fixed set of individuals) justice is concerned with the patterns of distributions and not with what particular individuals get. The distribution $\langle 2, 1 \rangle$ (two to the first person and 1 to the second) is

equally as just as $\langle 1, 2 \rangle$. There are, in general, powerful reasons for rejecting Anonymity. One is that justice is sensitive to the past (e.g., what commitments individuals have made or what good or bad deeds they have committed), but Anonymity (at least in its typical crude formulation) makes such information irrelevant. A second problem with Anonymity is that it is incompatible with individuals having differential rights over things (e.g., over their own bodies!). Anonymity holds that if it is just for me to increase my well-being from 5 to 10 units by using my body in a certain way while leaving you with 5 units, then it must also be just for *you* to increase your well-being from 5 to 10 units by using my body, while leaving me with 5 units. Again, this is quite implausible.⁴

Although Anonymity in its crude form is implausible, it may well be plausible as a condition on the distribution of benefits for which agents are in no way accountable (e.g., have no entitlement or desert claims). It is, for example, arguably a plausible condition on the distribution of initial (or brute luck generated) non-personal resources. In what follows, I shall, as Sen typically does, limit my attention to cases where there are no differential claims to benefits (e.g., from rights or desert). In such contexts, Anonymity is much more plausible. Even here, however, I believe that Anonymity should be rejected, but once again it would take us too far astray to examine this issue. Hence, we shall assume it in what follows.

A third condition that we shall generally assume is:

Strong Pareto: (1) If each person has the same benefits in one distribution as she does in another, then the two distributions are equally just. (2) If each person has at least as much benefits in one distribution as she does in another, and at least one person has more benefits in the first distribution as in the second, then the first distribution is more just than the second.⁵

The first part of Strong Pareto holds that if, for each person, the benefits of one distribution are the same as those of a second, the two distributions are equally just. This is not plausible in general because justice depends at least in part on the extent to which the *will* (or choices) of people (as opposed to the benefits they get) is respected. The distribution in which I am healthy because I choose not to smoke is not equally just with the one in which I am equally healthy and have the same other overall benefits because you force me not to smoke (and compensate me for the use of force), and no one else is affected. The first part of Strong Pareto rules out adequate sensitivity to choice-protecting rights. In what follows, however, we'll focus on contexts in which there are no relevant rights (or desert claims) at issue. In such contexts, this part of Strong Pareto is plausible.

The second part of Strong Pareto is an efficiency condition. It says that justice is positively sensitive to the benefits people receive. If one distribution is *Pareto superior* to a second—that is, if it gives some people more benefits and no one less—then it is more just. Although this is a relatively weak efficiency requirement (much weaker than the requirement that total benefits be maximized), it has significant implications. It entails that distributive considerations (e.g., equality) never trump efficiency in the sense of Pareto superiority. A Pareto superior distribution is more just—no matter how bad it is from a distributive viewpoint. For example, $\langle 1, 99 \rangle$ is deemed more just than $\langle 1, 1 \rangle$. Distributive principles, according to this condition, are limited in application at most to assessing the relative justice of *Pareto incomparable* distributions (i.e., pairs of distributions for which one distribution gives some people more and gives other people less). One positive implication of this limitation is that *leveling down* (reducing the benefits of better-off people) always makes things less just. As such, it is highly plausible. Equality promotion, for example, may well be relevant to justice, but only as a way of making some people's lives better. Making the better-off worse off without

benefiting anyone makes things less just, not more just.

In general, then, Strong Pareto is, it seems, a plausible principle of justice. We shall return below to its assessment when we consider certain distributive principles.

The final condition that we shall generally assume appeals to the notion of a *non-reversing downward transfer*, which is a transfer of benefits from a better-off person to a that does not make the originally worse-off person better off than the originally better-off person (i.e., does not reverse their positions in terms of being better off; it can make them equally well off). The move from $\langle 8,4 \rangle$ to $\langle 6,6 \rangle$, for example, involves a non-reversing downward transfer of two units. The move from $\langle 8,4 \rangle$ to $\langle 5,7 \rangle$, however, is a reversing transfer. Consider then:

Pigou-Dalton: If one distribution can be obtained from another by a non-reversing downward transfer, then it is more just than the other is.⁶

Pigou-Dalton introduces a very weak kind of distribution sensitivity. It says roughly that increasing the benefits of those with fewer benefits by a given amount is morally more important than increasing the benefits of those with greater benefits by the same amount. More exactly, it says that transferring a given amount of benefits from a better-off person to a worse-off person makes things more just when this leaves the originally better-off person at least as well off as the originally worse-off person. For example, it says that $\langle 6,6 \rangle$ is more just than $\langle 4,8 \rangle$.

Pigou-Dalton is a very weak, and highly plausible, kind of distribution-sensitivity. It holds only where the total is not affected. Most theorists endorse much stronger principles of distribution-sensitivity, but we shall start with this one.

Below, we shall examine three distribution-sensitive kinds of principles of justice: sufficientarianism (favoring distributions that better ensure that each person has an adequate

amount), prioritarianism (favoring distributions that better benefit those who are worse off), and egalitarianism (favoring more equal distributions).⁷

3. Sufficientarianism

The sufficiency view of justice is concerned with ensuring, to the extent possible, that each person receive an *adequate* amount of the benefits.⁸ Obviously, this requires a criterion for how much is adequate. Typically, the criterion of adequacy is something like enough (either in general or for that specific person) to meet basic (e.g., physiological) needs, avoid poverty, or have a minimally decent life (e.g., be able to show oneself in public without shame). Following Sen, we shall not address this important question and simply assume that some particular level of benefits has been selected as the adequacy (poverty) line.

As a theory of justice, (pure) sufficientarianism holds:

Sufficientarianism: A distribution is at least as just as another is if and only if it involves less or equal insufficiency.

This, of course, doesn't tell us much until we know how insufficiency of a distribution is measured.

Sen essentially created the framework for poverty (and insufficiency) measurement.⁹ All measures of poverty, he claimed, should satisfy the following conditions:

Weak Scale Invariance: Multiplying the poverty line and all the benefit levels by the same positive number does not affect aggregate poverty.

This entails that the chosen unit of measurement (e.g., pennies or dollars) does not affect the measurement of aggregate poverty, which is highly plausible. It also entails, however, that aggregate poverty is *purely relative* to the poverty line, and this is controversial. It requires, for example, that doubling everyone's real benefits while also doubling the poverty line (e.g., because poverty is understood in relative terms) does not affect aggregate poverty. Some might argue, however, that this will reduce aggregate poverty on the ground, for example, that there is less aggregate poverty when everyone has a \$10 shortfall from a poverty line of \$100 than when everyone has a \$5 shortfall from a poverty line of \$50—since in the former case each person is better off in absolute terms.¹⁰ Weak Scale Invariance, however, requires that aggregate poverty be unaffected. For simplicity, I shall ignore this objection, and accept Weak Scale Invariance.

A second condition imposed by Sen is:

Focus: The assessment of poverty (insufficiency) is not affected by changes in benefits above the poverty (adequacy) line, when those changes do not change the poverty line.

This simply ensures that the measure is a measure of poverty. For any fixed poverty line, how well off the non-poor are is irrelevant to aggregate poverty.¹¹ Following Sen, we shall focus on the measurement of poverty for a given poverty line.

Focus is indeed a plausible condition for the assessment of poverty (insufficiency), but it is precisely this condition that casts doubt on sufficientarianism as theory of justice. It ensures that both Strong Pareto and Pigou-Dalton are violated. Strong Pareto is violated because Focus requires that increasing the benefits of a *non-poor* person not reduce poverty (insufficiency). Sufficientarianism then concludes that such an increase does not make the distribution more just, which seems implausible. Benefits to those below the poverty line matter more for justice than

benefits to those above, but—as claimed by Strong Pareto—additional benefits to those above are better, from the viewpoint of justice, than no additional benefits at all. Pigou-Dalton is also violated, since Focus requires that a non-reversing downward transfer from a better-off *non-poor* person to a worse off *non-poor* person does not make things better from the viewpoint of eliminating poverty. Sufficiency then concludes that such a transfer does not make the distribution more just, which violates Pigou-Dalton and seems implausible. Such a transfer at least sometimes increases the equality of the distribution of benefits, and this seems to make things more just.

Because of Focus, sufficiency is arguably insufficiently demanding.¹² It finds both radical inequality and radical inefficiency morally acceptable as long as everyone has enough. This may be plausible if the benefits to be distributed are benefits the differential production of which is attributable to the choices of the agents (and not due to differential brute luck in capacities, opportunities, etc.). If, however, the benefits to be distributed are, as we are assuming, only those acquired by brute luck (i.e., due to factors over which the agent had no deliberative influence; e.g., manna from heaven), then the sufficiency view is, I would argue, insufficiently demanding.

Let us continue nonetheless to investigate how aggregate poverty (insufficiency) might be measured in light of the above conditions. One of the simplest measures is the *head count measure*, which simply counts the number of people below the poverty line. A major problem with this measure is that it holds that aggregate poverty is not affected by giving benefits to poor people that do not make them non-poor. Improving a person's situation from abject poverty to slight poverty does not change the number of poor people, and is thus deemed, on this measure, not to affect aggregate poverty. This is implausible.

More generally, any plausible measure of aggregate poverty, Sen rightly holds, should

satisfy the following condition:

Monotonicity (Below the Poverty Line): Increasing a *poor* person's benefits reduces aggregate poverty (insufficiency).

This is a kind of restricted Pareto condition applied to poverty measurement rather than justice. Although an increase in a *non-poor* person's benefits does not affect aggregate poverty (as required by Focus), increasing a *poor* person's benefits does, and it (of course) reduces aggregate poverty. This is a highly plausible condition, and the implausible head count measure violates it.

Consider now a second simple measure of aggregate poverty. The *total gap measure* simply adds up each person's shortfall from the poverty line (where a person's shortfall is zero, if she is at or above the poverty line). This satisfies Monotonicity, but it suffers from a different problem. Aggregate poverty is surely reduced if one non-reversingly transfers certain benefits from one poor person to a worse off poor person. The total gap measure, however, fails to generate this judgement. It is insensitive to the distribution of benefits among the poor. It measures only the total shortfall. A distribution-sensitive measure of aggregate poverty is needed.

More generally, Sen rightly imposes the following condition:

Weak Transfer Axiom: Aggregate poverty (insufficiency) is reduced by a non-reversing transfer of benefits from a better-off *poor* person to a worse off *poor* person.

This is just Pigou-Dalton restricted to transfers below the poverty line. Although downward non-reversing transfers of benefits that occur *above* the poverty line do not affect aggregate poverty

(i.e., unrestricted Pigou-Dalton applied to poverty measurement is violated, as required by Focus), such transfers *below* the poverty line reduce poverty. The total gap measure, however, violates this condition (since the total gap is unaffected).

Sen proposes the following measure, S , of aggregate poverty in a given distribution y :

$$S(y) = [HCR(y) \cdot PGR(y)] + [HCR(y) \cdot (1 - PGR(y))] \cdot G_p(y)$$

where: $HCR(y)$ is the head count ratio for y (i.e., number of poor divided by population size);

$PGR(y)$ is the poverty gap ratio for y (i.e., total poverty gap divided by the product of the poverty line and the number of the poor);

$G_p(y)$ is the inequality in the distribution of benefits among the poor in y as measured by the Gini coefficient (a standard measure of inequality).

Sen (1976a) proves that where large numbers of the poor are involved, roughly this measure follows from two assumptions: (1) For a fixed population size, aggregate poverty is ranked on the basis of the *sum* of the *weighted* poverty gaps, where the weight of a given gap is equal to its ordinal position in the size of such gaps (i.e., if there are n people with poverty gaps, the largest gap has a weight of n , the second largest gap has a weight of $n-1$, and the smallest gap has a weight of one). (2) To handle variable population size, the sum of the weighted poverty gaps is normalized by dividing by the population size.¹³

Each of these assumptions is, as Sen clearly recognizes, controversial. The assumption that aggregate poverty is the sum of the weighted poverty gaps, with the above ordinal position weights, can be challenged on at least three grounds. One is that aggregate poverty might not be ranked on the basis of the sum of weighted poverty gaps. Consider, for example, leximin poverty

gap. It holds that there is at least as much aggregate poverty in one distribution as in another if and only if the *largest* poverty gap in the first is at least as great as that in the second, and, if they are equally great, the second largest poverty gap in the first is at least as great as that in the second, and so on. This measure satisfies Weak Scale Invariance, Focus, Monotonicity, and Weak Transfer Axiom, and, as well known, it is not based on the sum of (finitely) weighted poverty gaps.¹⁴ Leximin poverty gap is not, however, a very plausible measure of aggregate poverty (because of the absolute priority that it gives to the worse off). Although there are other possibilities, it seems reasonable to grant Sen the assumption that aggregate poverty is ranked on the basis of the sum of (finitely) weighted poverty gaps.

A second challenge to Sen's first assumption is that the particular weights that Sen assumes for weighting poverty gaps seem quite arbitrary. The weights should indeed be greater for greater poverty gaps (as required by Weak Transfer Axiom), but it's not at all clear why they should correspond precisely to the ordinal position of the poverty gap. More specifically, this weighting scheme seems inappropriately insensitive to the differences in *magnitude* of the poverty gaps beyond those differences reflected in the ordinal rank. For example, suppose that there are only two people in society, the poverty line is 10, and both are poor in the distributions that we consider. Sen's weighting scheme assigns a weight of 2 to the larger poverty gap and a weight of 1 to the smaller poverty gap—no matter what the size of these poverty gaps. Thus, a weight of 2 is given to the poverty gap of 9 by the first person in $\langle 1, 9 \rangle$ and is also given to the poverty gap of 2 to the first person in $\langle 8, 9 \rangle$. The fact that the former poverty gap is much larger is deemed irrelevant. Only the ordinal position of the poverty gaps matters for the weights. This seems implausible.

A third challenge to Sen's first assumption is that, as he recognizes, it violates the following condition:

Strong Transfer Axiom: Aggregate poverty (insufficiency) is reduced by a non-reversing transfer of benefits from a better-off (poor or non-poor) person to a worse off *poor* person.

This is like the above Weak Transfer Axiom, except that it does not require that the donor be poor. Given that Sen weights poverty gaps by their ordinal poverty position, his poverty measure violates this condition for certain transfers that move the donor from being non-poor to being poor. Such transfers increase the number of poor people and this can have the result of increasing Sen's measure of poverty. (This is so, for example, when there are just two people, and the better-off person is initially just barely above the poverty line.) It seems fairly plausible, however, that sufficientarian justice is increased by such a transfer. Even though the number of poor people increases, the total poverty gap does not increase, and the poverty gaps are spread more equally. Thus, Strong Transfer Axiom seems plausible, and yet it is violated by Sen's assumption about weighted poverty gaps.¹⁵

Sen's first assumption—that aggregate poverty is based on the sum, suitably normalized for population size, of the weighted poverty gaps, with his ordinal position weighting scheme—thus is implausible for a sufficientarian theory of justice. His second condition—that a suitable normalization for population size is to divide the sum by the population size—is, I shall argue, also implausible.

The general question here is that of whether poverty (sufficiency) is to be understood in (as Sen proposes) population-relative (i.e., per capita) terms or absolute terms. Where 10 is below the poverty line, $\langle 10, 100, 10, 100 \rangle$ has twice as much absolute poverty as $\langle 10, 100 \rangle$ but has the same per capita poverty.

Fortunately, in the present context, we need not resolve this general issue. In the context

of Sen's total ordinally weighted poverty gap view, normalizing by dividing by the number of people (i.e., relativization to population size) implausibly has the effect that adding non-poor people to the population reduces aggregate poverty—since it increases the denominator of his measure (the population size) without increasing the numerator (the sum of the weighted poverty gaps). This, I claim, is an implausible result for a sufficientarian theory of justice.

Another way of stating the above concern is that aggregate poverty should satisfy a strong version of a “focus” (on the poor) condition. Above we noted that Sen's measure satisfies Focus, which states that aggregate poverty is not affected by increasing the *benefits* of the non-poor. The following stronger focus condition requires in addition that aggregate poverty not depend on the *number* of the non-poor.

Strong Focus: The assessment of poverty (insufficiency) is not affected by changes in benefits to those above the poverty (adequacy) line, *nor by the number of non-poor people*, when this does not change the poverty line.

Aggregate poverty, that is, should be only about the condition of the poor and not depend in any way on the condition (number or benefit levels) of the non-poor.

This concludes my discussion of Sen's work on the measurement of poverty.

Philosophers have recently begun to give more attention to sufficientarian theories of justice, but few have systematically investigated how aggregate sufficiency is to be measured. When it is feasible to give everyone sufficient amounts of the relevant benefits, there is no need for a theoretical measure of aggregate sufficiency (since any adequate measure will agree that there must be perfect sufficiency). Given, however, that in many contexts it may not be possible to give everyone a sufficient amount, the development of a measure of aggregate sufficiency is

indeed important. The work of Sen and other economists thus provides an important framework for the investigation of this issue.¹⁶ I have suggested, however, that neither the rank order weighting system invoked by Sen's measure of poverty, nor his relativization to population size are plausible for a sufficientarian theory of justice. This is not a criticism of Sen's work, since he did not intend his measure to be so used.

4. Prioritarianism

The sufficiency view of justice is concerned with increasing the benefits that people have, but only up to an adequate amount. It gives absolute priority to a benefit, no matter how small, to someone whose benefits will remain inadequate, over a benefit, no matter how large, to a person whose benefits are already adequate. Prioritarianism generalizes this idea. It holds that increasing the benefits that people have always matters (and not just below the adequacy level), but a given increase matters more morally when it is given to a person who has a lower initial level of benefits than it does when it is given to someone who has a higher initial level. There is, it is claimed, a kind of decreasing marginal moral importance to marginal increases in benefits.¹⁷

Because prioritarianism holds that increasing a person's benefits always makes things more just, it (unlike sufficientarianism) satisfies Strong Pareto (and not merely Monotonicity). Furthermore, because it holds that a given benefit increase is always more important when it occurs at a lower initial of benefits, it (unlike sufficientarianism) satisfies Pigou-Dalton (and not merely Weak Pigou Dalton). This, I believe, makes it more promising as a theory of justice.

There are several different forms that prioritarianism can take, but the following are the three main ones:

Leximin: One distribution is at least as just as a second if and only if (1) the person with the

least benefits in the first distribution has at least as much benefits as the person with the least benefits in the second distribution, and (2) if there is a tie in the previous comparison, then the person with the *second* least benefits in the first distribution has at least as much benefits as the person with the *second* least benefits in the second distribution, and so on for the persons with the third, etc. least benefits.

Strong Prioritarian Totalism : One distribution is at least as just as a second is if and only if its total weighted benefits are at least as great, where the weights for a given increment finitely decrease as the benefits prior to the increment finitely increase.

Weak Prioritarian Totalism: One distribution is at least as just as a second is if and only if (1) it has greater total benefits, or (2) it has the same total benefits and is at least as just according to leximin.

Leximin and weak prioritarian totalism are the polar extremes of prioritarianism. Leximin assigns infinitely greater weight to those who are worse off. It favors a benefit, no matter how small, to a person who remains a worst-off person over a benefit, no matter how large, to a better-off person. Weak prioritarian totalism (of which prioritarian utilitarianism is an example), by contrast, assigns only infinitesimally greater weight to those who are worse off. Where the total benefits are not equal, weak prioritarian totalism agrees with standard totalism, but, where the benefits are equal, it agrees with leximin. Given that, on most conceptions of benefits, the total benefits will typically not be equal, weak prioritarian totalism typically gives a very limited role to leximin. The role it gives, however, is sufficient (just barely!) to ensure that Pigou-Dalton is satisfied. It thus has a claim (just barely) to be a prioritarian theory.

Between these extreme forms of prioritarianism is strong prioritarian totalism . It is like totalism (e.g., total utilitarianism) except that it adds up *weighted* benefits, where the weight assigned to a one-unit increment to a certain level of benefits is *finitely* greater—and not infinitely or infinitesimally greater—than the weight assigned to a one-unit increment to any *higher* level. Thus, for example, the increment from 0 to 1 might have a weight of 1, the increment from 1 to 2 might have a weight of 1/2, and an increment from 2 to 3 might have a weight of 1/3. A person with three units of benefits generates 1 5/6 units of weighted benefits, whereas three people each with one unit generates 3 units of weighted benefits. The decreasing marginal weights generate extra priority for those who are worse off, but, given that the extra weight is only finite, a sufficiently larger benefit to a better-off person can produce a greater increase in weighted benefits than a given benefit to a worse-off person.

Sen contributed much to our understanding of leximin and weak prioritarian totalism (as well as standard totalism).¹⁸ The following material will bring together some of the results of Sen and others on the topic of prioritarianism.

Leximin, weak prioritarian totalism, and strong prioritarian totalism each satisfy Ordering, Anonymity (our impartiality condition), Strong Pareto (our efficiency condition), and Pigou-Dalton (our distribution-sensitivity condition). They each also satisfy the following condition, which, by definition, is satisfied by all prioritarian theories:

Strong Separability (with respect to unaffected individuals): The justice ranking of two distributions, x and y , is fully determined by the benefits of the affected individuals (i.e., the individuals whose benefits are not the same in each distribution) and does not depend on the benefits of the unaffected individuals (i.e., those whose benefits are the same in each distribution).

This condition says that if some set of individuals is unaffected by the choice between x and y (each gets the same benefits in x as in y), then the ranking of x and y does not depend on what their benefits are. For example, the ranking of $\langle 1,4,6,7 \rangle$ versus $\langle 2,4,3,7 \rangle$ (for which the second and fourth persons are unaffected) must be the same as the ranking of $\langle 1,9,6,2 \rangle$ versus $\langle 2,9,3,2 \rangle$ (for which the second and fourth persons are unaffected). With respect to affected individuals (the first and third persons), the first pair of distributions is identical to the second pair: $\langle 1,-,6,- \rangle$ vs. $\langle 2,-,3,- \rangle$ in both cases. Strong separability requires that the first pair be ranked the same way as the second pair. The particular values for the second and fourth persons are deemed irrelevant (given that they are unaffected).

Strong Separability is not uncontroversial. Various forms of egalitarian theory, for example, violate it. All prioritarian views (as well as totalism), however, satisfy it by definition (since they always favor making people better off no matter what the situation of those who are unaffected), and so we shall temporarily grant it.¹⁹

Let us now consider some conditions that distinguish Leximin, strong prioritarian totalism, and weak prioritarian totalism from each other. Consider first:

Irrelevance of the Cardinality of Benefits: For any two distributions, transforming each person's benefits by a positive monotonic transformation, the same for all persons, does not affect their relative ranking with respect to justice (if x is at least as just as y before the transformation, then it is also so after the transformation).

A positive monotonic transformation of benefits is a transformation that preserves the relative size of benefits (equal benefits are transformed to equal benefits and larger benefits are

transformed to larger benefits). For example, $\langle 1,2,9 \rangle$ is a positive monotonic transformation of $\langle 2,4,6 \rangle$, but $\langle 3,2,5 \rangle$ is not. This condition says that the numbers representing benefits are irrelevant except, perhaps, for the ordinal information they represent (for a given person and perhaps also between people). It does not rule out interpersonal comparability of benefit levels, but it rules out any appeal to cardinality of benefits, and a fortiori any interpersonal comparisons of zero levels or of units.

One way of defending leximin is the following:

Observation 1: If the justice ranking relation satisfies Ordering, Anonymity, Strong Pareto, Pigou-Dalton, Strong Separability, and Irrelevance of Cardinality, then leximin is the justice ranking relation.²⁰

In the context of prioritarian theories, the only controversial condition here is Irrelevance of Cardinality. If benefits are only ordinally measurable (with assessments that one benefit is equal to or greater than another, but with no assessment of how much greater one benefit is), then, of course, Irrelevance of Cardinality is plausible. Here, however, we are abstracting from the specifics of any given conception of benefits, and thus we cannot answer the question of whether the benefits are merely ordinally measurable.

There is, however, a further issue, even if benefits are cardinally measurable: that cardinality may be irrelevant to justice. Justice may depend solely on ordinal information about benefits. If so, then Irrelevance of Cardinality is plausible. Nonetheless, it seems highly plausible that if we can assess, for a given person, not merely that one distribution is better or worse for her, but also how much better or worse it is, then surely this is relevant to justice. Hence, Irrelevance of Cardinality is not a very plausible condition, if we assume a conception of benefits

for which there is some cardinality. Thus, this condition does not in general (independently of a specific conception of benefits) provide a justification for leximin.

Consider, then, the following condition:

Hammond Equity: If two distributions give everyone, except two individuals, the same benefits, and one of these two individuals has less benefits than the other person under one distribution, and less or equal benefits under the other distribution, then the distribution that gives this worse-off person more benefits is more just than the other distribution.

Consider, for example, $\langle 1, 5, 10 \rangle$ and $\langle 2, 3, 10 \rangle$. The only difference between these two distributions concerns the benefits that the first and second person get. The first person gets less than the second person does in both distributions (1 vs. 5 in the first distribution, and 2 vs. 3 in the second). Hammond Equity therefore requires that the second distribution (which is better for the worse-off person) be judged more just. Both weak prioritarian totalism and strong prioritarian totalism violate this condition (since both allow that if the loss to the better-off person is sufficiently greater than the benefit to the worse-off person, then the result is less just). Leximin satisfies it. Indeed, an important result is:

Observation 2: If the justice ranking relation satisfies Ordering, Anonymity, Strong Pareto, and Hammond Equity, then leximin is the justice ranking relation.²¹

This is an important result for *understanding* leximin, but it fails, I believe, to offer a *justification* of leximin. Hammond Equity is, I claim, implausible. To see this, consider $\langle 1, 99, 100, \dots, 100 \rangle$ and $\langle 2, 3, 100, \dots, 100 \rangle$, where the average is slightly less than 100 and where

the difference between 3 and 100 units is very significant. Hammond Equity requires that the second distribution be judged more just, since it is better for the worst off person. This, I claim, is implausible. Benefits to the worse-off person do not have such absolute priority. In particular, when, for two distributions, (1) the two affected individuals are both below average in both, (2) compared with the second distribution, the first distribution is only *slightly* better for the worse-off person but *very significantly* worse for the better-off person, and (3) the first distribution leaves the worse-off person *only slightly* worse off than the better-off person, then the second distribution is more (not less) just.

It's worth noting that Irrelevance of Cardinality (assuming that benefits are cardinally measurable) and Hammond Equity are, I suggest (without argument), not merely conditions that need not be satisfied, but conditions that a theory *must violate* in order to be plausible. Any theory that makes available information about cardinality irrelevant, or that always gives absolute priority to a worse-off person, is implausible. Given that leximin satisfies these two conditions, we have, I suggest, reason for rejecting leximin. Again, obviously this is controversial and requires a more careful examination than that given here.

Any plausible theory, I have suggested, will violate Hammond Equity. This thought is roughly captured by the following condition:

Minimal Aggregative Efficiency: At least sometimes, decreasing the benefits of one person and increasing the benefits of a better-off person by a greater amount makes things more just.²²

This is essentially just the denial that benefits to a worse-off person always have absolute priority. As such, it is extremely weak, and likely to be accepted by a wide range of theorists. Leximin, however, violates this condition.

Let us now turn to Weak Prioritarian Totalism, and consider the following condition:

Irrelevance of External Reference Points: For any two distributions, modifying everyone's benefits by multiplying by a positive constant (perhaps one) and then adding a constant (perhaps zero), the same two constants for all persons, does not affect the ranking of the two distributions with respect to justice (if x is at least as just as y before the modification, then it is also so afterwards).²³

The idea here is that there are no external reference points—for example, natural zero for benefits, level of adequacy for benefits, or upper or lower bounds on benefits—that are relevant to justice. According to this condition, justice is a purely relative matter, and changing everyone's benefits in the same specified way produces a distribution that is equally just with the original. If some external reference point were relevant, then at least sometimes adding a constant benefit to everyone would, for example, make a difference to the justice ranking by changing someone's relationship to the reference point in a relevant way (e.g., moving someone from below to above the adequacy level). The above condition holds, for example, that the ranking of $\langle 2,5 \rangle$ and $\langle 1,6 \rangle$ must be the same as that of $\langle 5,8 \rangle$ and $\langle 4,9 \rangle$ (the latter pair is obtained from the former pair by adding three to each person's benefits).

The following result is worth noting:

Observation 3: If the justice ranking relation satisfies Ordering, Anonymity, Strong Pareto, Strong Separability, Pigou-Dalton, Minimal Aggregative Efficiency, and Irrelevance of External Reference Points, then weak prioritarian totalism is the justice ranking relation.²⁴

The only theories satisfying Ordering, Anonymity, Strong Pareto, and Strong Separability are leximin, leximax (making the best-off person as well off as possible, etc.), strong prioritarian totalism and related views, totalism, and Weak Prioritarian Totalism. Pigou-Dalton rules out totalism and leximax. Minimal Aggregative Efficiency rules out leximin. Irrelevance of External Reference Points rules out (as I'll explain below) strong prioritarian totalism and related views. Weak prioritarian totalism is thus the only theory satisfying all these conditions.

The crucial question here is whether Irrelevance of External Reference Points is plausible. Leximin satisfies this condition. It gives absolute priority to a worse-off person no matter what the benefit levels (e.g., it favors $\langle 2,5 \rangle$ over $\langle 1,6 \rangle$ and, for any n , favors $\langle 2+n,5+n \rangle$ over $\langle 1+n,6+n \rangle$). Weak prioritarian totalism also invokes no external reference point. If one distribution has a greater total than another does, this remains true if n units are added to everyone's benefits. Furthermore, if one distribution has the same total, but is leximin better (e.g., $\langle 2,5 \rangle$ vs. $\langle 1,6 \rangle$), this remains true if n units of benefit are added to everyone.

Strong prioritarian totalism, however, does require, except under very special circumstances, the relevance of an external reference point. It requires that the weights be anchored in some external reference point and not merely determined in relative terms, which would result in a violation of Strong Separability. More specifically, it requires that for any given interpersonally valid scale for measuring benefits (no matter how the zero point and unit are set), there is a corresponding decreasing positive weighting scale for benefits. If the scale for measuring benefits is changed (by changing the zero point or the unit), the weighting scale must also be correspondingly changed. For example, if the weight for the increase from 4 to 5 units of benefit on a given scale is n , and then the scale is changed by doubling the size of the units, then the weight on this new scale for the increase from 2 (which was 4, on the old scale) to 2.5 (which was 5, on the old scale) must also be n .

Irrelevance of External Reference Points is not, I claim, a very plausible condition. On any plausible conception of benefits, it will be possible to distinguish, for example, cases where a worse-off person is in abject poverty (relative to those benefits) and cases where she is relatively prosperous (both of which appeal to an external reference point; benefits are not viewed merely in relative terms). More specifically, it's plausible that the extra weight given by justice to the benefits of a worse-off person compared with those of a person who is n units better off is greater when the worse-off person is abjectly poor than when she is very affluent. Irrelevance of External Reference Points, however, rules out these appeals to absolute poverty and affluence, and is thus not very plausible.

We should then, I suggest, reject Irrelevance of External Reference Points. Indeed, a theory of justice is plausible only if it violates this condition. Thus, we have some reason for rejecting weak prioritarian totalism (as well as totalism) and leximin.

Let us turn now to strong prioritarian totalism, and consider the following condition:

Continuity: For any two distributions, x and y , if there is an infinite sequence of distributions such that (1) the sequence converges (at the limit) to x (i.e., for each person, the benefits to her in the sequence converge to her benefits in x), and (2) each distribution in the sequence is at least as just as y , then x is at least as just as y .

The core idea here is twofold: (1) If all the distributions in the sequence that converges to x are *equally as just* as y , then x should also be judged equally as just. (2) If all the distributions in the sequence that converges to x are *more just* than y , then x should either also be judged as more just than y , or it should be judged as equally as just. Distribution x should not be judged as less just, since this would involve a kind of gap (x is less just but all the members of the

sequence that converges to x are more just).

We can note the following conjecture (suggested to me by Bertil Tungodden):

Observation 4: If the justice ranking relation satisfies Ordering, Anonymity, Strong Pareto, Pigou-Dalton, Strong Separability, and Continuity, then strong prioritarian totalism is the justice ranking relation.²⁵

Continuity is violated by leximin and Weak Prioritarian Totalism, but is satisfied by strong prioritarian totalism. To see that leximin violates Continuity, consider a two person case with an infinite sequence of distributions starting with $\langle 2, 2 \rangle$ that approaches (but never reaches) $\langle 1, 1 \rangle$ at the limit (e.g., $\langle 2, 2 \rangle$, $\langle 1.5, 1.5 \rangle$, $\langle 1.25, 1.25 \rangle$, etc.). Each member of this sequence will be judged better off than $\langle 1, 3 \rangle$ by leximin (because the first person is the worst off and is better off), but $\langle 1, 1 \rangle$ is judged worse than $\langle 1, 3 \rangle$ (because both people are worst off persons and each is worse off). This violates Continuity. To see that weak prioritarian totalism violates continuity, consider an infinite sequence of distributions starting with $\langle 3, 1 \rangle$ that approaches (but never reaches) $\langle 2, 0 \rangle$ at the limit (e.g., $\langle 3, 1 \rangle$, $\langle 2.5, .5 \rangle$, $\langle 2.25, .25 \rangle$, etc.). Each member of this sequence will be judged better off than $\langle 1, 1 \rangle$ by weak prioritarian totalism (because each member has a greater total), but $\langle 2, 0 \rangle$ is judged worse off (because it has the same total, but is leximin inferior). This violates Continuity. Strong prioritarian totalism, on the other hand, satisfies Continuity, roughly because it gives only finitely (and not infinitesimally or infinitely) extra weight to the worse off.

Strong prioritarian totalism, it turns out, is the *only* theory satisfying Continuity and the conditions above. Continuity thus provides support, in the prioritarian context, for strong prioritarian totalism. It's doubtful, however, that we should accept Continuity as a condition on

justice. It would be formally nice for justice to behave in a continuous fashion, but it's far from clear that justice must do so. It may, for example, be implausible for leximin to give absolute priority to the worse off, but this implausibility stems from the *substantive* implausibility of such a judgement and not, it seems, from the formal requirement of Continuity. Continuity rules out treating some considerations as lexically posterior to others, but there is nothing incoherent, or even obviously implausible, about such treatment.

Let us consider one final condition that supports strong prioritarian totalism :

Extended Pigou-Dalton: At least sometimes, decreasing one person's benefits and increasing a worse-off person's benefits by less, when this still leaves the originally worse-off person no better off than the originally better-off person, makes things more just.

Like Pigou-Dalton, this requires that, under certain conditions, justice be increased by certain non-reversing downward transfers. Unlike Pigou-Dalton, which deals with "even" ("fixed amount") downward transfers, Extended Pigou-Dalton deals with "inefficient" downward transfers, which are downward transfers where the benefit to the recipient is less than the cost to the donor. Some benefits are lost. Furthermore, unlike Pigou-Dalton, which holds that "even" non-reversing, downward transfers *always* increase justice, Extended Pigou-Dalton only requires that "inefficient" non-reversing, downward transfers *sometimes* increase justice.

Extended Pigou-Dalton is a very weak condition, and is likely to be endorsed by a wide range of theorists. It is satisfied by leximin and by strong prioritarian totalism , but it is violated by Weak Prioritarian Totalism. The latter always favors a distribution with a greater total, and thus (like standard totalism) never favors inefficient transfers, whereas Extended Pigou-Dalton requires that sometimes inefficient, non-reversing downward transfers make things more just.

Consider, then:

Observation 5: Strong prioritarian totalism satisfies Ordering, Anonymity, Strong Pareto, Pigou-Dalton, Strong Separability, Minimal Aggregative Efficiency, and Extended Pigou-Dalton, but leximin violates Minimal Aggregative Efficiency and weak prioritarian totalism violates Extended Pigou-Dalton.

Imposing Minimal Aggregative Efficiency and Extended Pigou-Dalton is thus sufficient to rule out Leximin and Weak Prioritarian Totalism, and leave standing strong prioritarian totalism. The above conditions, however, do not fully characterize the latter view, for there are other related theories that satisfy all these conditions (see example in note). Thus, adding Extended Pigou-Dalton brings us closer to fully characterizing strong prioritarian totalism but does not suffice. I do not know whether anyone has fully characterized strong prioritarian totalism without appealing to Continuity.²⁶

Strong prioritarian totalism is, I believe, the most plausible prioritarian view. Even it, however, is subject to a powerful objection—at least where the benefits being distributed are generated by brute luck. Suppose that there are just two people: a person with no benefits and a person with enormous levels of benefits. Suppose that you can give either a large benefit to the rich person or a much smaller benefit to the poor person. The Finitely Weighted Total View says that, if the larger benefit to the rich person is sufficiently larger, then it is more just to give that benefit to the rich person than to give the much smaller benefit to the poor person. This is because, no matter how finitely much greater weight is given to the benefits of the poor person, the increased weighted benefits for the rich person will be greater, if the increase in benefits is sufficiently greater. This seems, however, to be mistaken. In a context in which everyone has an

equal claim to benefits (e.g., the distribution of brute luck benefits), justice gives *absolute priority* to benefits to a person who remains with no more than an average share over those to individuals who already have at least an average share. Leximin is at least right about this kind of case.

This objection gives a special role to the benchmark of an average benefit, and it will be raised by certain kinds of egalitarian theorists. Let us therefore now consider egalitarianism, our third distribution-sensitive theory of justice.

5. Egalitarianism

Pure egalitarian theories—which make justice depend solely on how equal the distribution of benefits is—satisfy Ordering, Anonymity, and Pigou-Dalton, but violate Strong Pareto. They violate the latter because they view perfect equality as more just than a Pareto-improvement thereto (e.g., $\langle 2, 2 \rangle$ is more just than $\langle 8, 9 \rangle$). This reflects the fact that pure egalitarianism is not in any way concerned with increasing the benefits that people receive. It is only concerned with how equally they are distributed. Following Sen, I have suggested that a plausible theory of justice (understood as concerned with what we owe others) will be both distribution-sensitive (satisfy at least Pigou-Dalton) and sensitive to some kind of efficiency consideration (satisfy at least Strong Pareto). Hence, pure egalitarianism is not, I believe, a plausible theory of justice.

In what follows, we shall be concerned with impure forms of egalitarianism. More specifically, we shall assume that a plausible version of egalitarianism may only satisfy the following weak egalitarian condition. The condition appeals to the notion of *anonymous Pareto incomparability*, which holds between two distributions if and only if the first distribution, *and each of its permutations* (i.e., same pattern of distribution but perhaps with individuals occupying different positions), is Pareto incomparable with the second (i.e., better for some people but

worse for others). For example, $\langle 3,5 \rangle$ is Pareto incomparable with $\langle 4,3 \rangle$, but it is not anonymously Pareto incomparable, since $\langle 3,4 \rangle$ is a permutation of $\langle 4,3 \rangle$ and is Pareto inferior to $\langle 3,5 \rangle$. Distribution $\langle 3,5 \rangle$ is, however, anonymously Pareto incomparable with $\langle 6,2 \rangle$, since it and its only distinct permutation, $\langle 5,3 \rangle$, are each Pareto incomparable with $\langle 6,2 \rangle$. Anonymous Pareto incomparability entails Pareto incomparability, but not vice-versa.

Consider then:

Weak Egalitarianism: If two distributions are anonymously Pareto incomparable, and one is more equal than the other is, then it is more just than the other is.

This is a very weak egalitarian condition. Not only is it silent when one distribution is Pareto superior to the other, it is also silent when one distribution is Pareto incomparable to the other but is not anonymously Pareto incomparable. Some egalitarians may want equality to be determinative whenever distributions are Pareto incomparable (even if not anonymously Pareto incomparable), but we shall start with this weak condition.

Weak Egalitarianism requires justice to be based partly on the promotion of equality, but it is compatible with Strong Pareto and our other background conditions (Ordering, Anonymity, and Pigou Dalton). Its content, however, is not clear, until we know how equality is measured. The rest of this section will be primarily concerned with the measurement of equality, a topic on which Sen has been extremely influential.²⁷

The concept of equality also satisfies Ordering (or at least reflexivity and transitivity), Anonymity, and Pigou-Dalton, as defined in the background section, but now understood as applying to the relation of being-at-least-as-*equal* rather than to the relation of being-at-least-as-*just*.

An obvious condition on equality is:

Perfect Equality: Perfect equality obtains when everyone has the same benefit level.

There are other conditions that the concept of equality satisfies, but these will suffice for our purposes. These conditions, as Sen and others have noted, entail that, where the population-size and mean are fixed, *Lorenz-domination*, as defined below, is sufficient for being more equal. One distribution Lorenz-dominates another if and only if (1) for any real number, n , inclusively between 0 and 100, the percentage of total benefits allocated to the poorest (in terms of the benefits) $n\%$ of the population is at least as great for the first distribution as it is for the second, and (2) for some real number, n , between 0 and 100, that percentage is greater. For example, for $\langle 1,2,3,4 \rangle$, the Lorenz values for 25%, 50%, 75%, and 100% of the population are 10%, 30%, 60%, and 100% respectively, and for $\langle 2,2,2,4 \rangle$ the corresponding values are 20%, 40%, 60%, and 100% respectively. Thus, the second distribution Lorenz-dominates the first. (For example, it gives 20%, rather than 10%, of the total benefits to the poorest 25%.) By contrast, for $\langle 1,3,3,3 \rangle$ the corresponding Lorenz values are 10%, 40%, 70%, and 100% respectively. It neither Lorenz-dominates, nor is Lorenz-dominated by, $\langle 2,2,2,4 \rangle$, since it gives a smaller share of the benefits (10% vs. 20%) to the poorest 25%, but gives a larger share (70% vs. 60%) to the poorest 75%.

The interesting result, connecting the concept of equality and Lorenz-domination, is this:

Observation 1: Given that the equality relation satisfies Ordering (more minimally: transitivity), Anonymity, and Pigou-Dalton, for any two distributions having the same population size and the same mean (i.e., average), if one distribution Lorenz-dominates the other, then it is more equal.²⁸

For a fixed population size and fixed mean benefit, Lorenz-domination provides the core

of the equality relation. This leaves open three issues for the measurement of equality for egalitarian theories of justice: (1) For a fixed population size and fixed mean benefits, what determines whether one distribution is more equal than another when neither Lorenz-dominates the other? (2) How are judgements of equality affected when the size of the mean benefit is varied? (3) How are judgements of equality affected when the population size is varied? Because of space limitations, I shall address only the first two questions.

With respect to extension of the equality relation beyond Lorenz-domination, for a fixed population-size and mean, a common condition is the following:

Diminishing Transfers: Equality is increased by the combination of (1) a strictly non-reversing downward transfer of m units ($m > 0$) from a person j to a person who has n units fewer of benefits ($n > 2m$), and (2) a strictly non-reversing upward transfer of m units from a person k , who is more than m units better off than j , to a person who has n units more benefits than k .

The idea of this principle is that the impact on equality of a non-reversing transfer of a given amount to a person whose benefit level is lower by a given amount is greater when it is between people with lower levels of benefits (e.g., among poor people) than when it is between people with higher levels of benefits (e.g., among rich people). There is, that is, a decreasing marginal impact on equality for a fixed transfer amount and a fixed transfer distance. Thus, for a fixed transfer amount and distance, if a downward transfer is combined with an upward transfer, the net result should increase equality as long as the transferor of the upward transfer is better off after the transfer than the transferee of the downward transfer was before the transfer.

This principle has a fair amount of plausibility. I believe, however, that it is slightly too strong. For it requires that $\langle 10, 10, 30, 70, 100 \rangle$ (mean of 44) be judged as *more equal* than

$\langle 0, 20, 40, 60, 100 \rangle$ (mean of 44)—even though it gives fewer benefits to those below the mean (50 vs. 60). The former distribution is obtainable from the latter by transferring 10 units from the person with 20 to the person with 0, and transferring 10 units from the person with 40 (slightly below the mean of 44) to the person with 60 (already above the mean). Diminishing Transfers does not take into account that some upward transfers are from people below the mean to people above the mean. Such transfers, I claim, can decrease equality when coupled with a non-reversing downward transfer of an equal amount and an equal distance. A slight weakening of Diminishing Transfers, however, is plausible:

Non-Upward-Crossing Diminishing Transfers: Equality is increased by the combination of (1) an even strictly non-reversing downward transfer of m units ($m > 0$) from a person j to a person who has n units fewer of benefits ($n > 2m$), and (2) an even strictly non-reversing upward transfer of m units from a person k , who is more than m units better off than j , to a person who has n units more benefits than k , where the transferor and the transferee of this upward transfer are on the same side of the mean (both above or both below).²⁹

This condition is exactly like the original except that it does not apply where the upward transferor is below the mean and the upward transferee is above the mean. Non-Upward-Crossing Diminishing Transfers is, I believe, a plausible principle. Even if it (or the original Diminishing Transfers) is accepted, that does not completely determine how equality is measured—even in the case of fixed population-size and mean. Sen discusses various measures, but space limitations prevent me from addressing this issue further.³⁰

Let us turn now to the main question of how changes in the mean benefit (for a fixed population size) affect equality. More specifically, we can ask what the impact on equality is of

(1) increasing everyone's benefit by the same amount, or (2) multiplying everyone's benefit by the same positive factor. The following are the main answers that have been given to this question.

Constant Additions Invariance: Equality is unaffected by increasing each person's benefits by the same amount.

This is the standard condition for an absolute conception of equality. It holds that equality is based on the absolute magnitude of the differences in people's benefits. Increasing everyone's benefits by the same amount does not affect these differences and thus does not affect the level of equality. Distribution $\langle 1,2 \rangle$ is equally equal with $\langle 2,3 \rangle$.

Constant Additions Improvements: Equality is increased by increasing each person's benefits by the same amount.

This is a standard condition for conceptions of equality that are at least partially relative to the mean. The idea here is that although constant additions do not affect the differences between individuals, they do make those differences a smaller percentage of the benefits that people have (since everyone has more). Hence, equality is increased.

Proportionate Increases Invariance: Equality is unaffected by multiplying each person's benefits by the same positive factor.³¹

This is the standard condition for a purely mean-relative view of benefits. It holds that equality is

based on the proportionate shares rather than the absolute benefit levels. Thus, increasing everyone's benefits by the same proportion does not affect the level of equality. Distribution $\langle 1,2 \rangle$ is equally equal with $\langle 2,4 \rangle$. In both cases, the first person has one third of the benefits.

Proportionate Increases Worsening: When some inequality is present, equality is decreased by multiplying each person's benefits by the same factor greater than one.

This is the standard condition for views that are at least partially absolute. It holds that $\langle 2,4 \rangle$ is less equal than $\langle 1,2 \rangle$, on the ground that in absolute terms there is a greater difference in benefits in $\langle 2,4 \rangle$.

For the purposes of egalitarian theories of justice, Proportionate Increases Invariance (and any purely relative view of equality) is, I would argue, implausible. There is much less relevant inequality in $\langle 1,2 \rangle$ than there is in $\langle 100,200 \rangle$. Indeed, Proportionate Increases Worsening seems highly plausible. The assessment of Constant Additions Invariance and Constant Additions Improvements is more difficult, and I here leave that issue open.

We have seen, then, that there are many controversial issues concerning the measurement of equality. Before concluding, we shall note a recent striking result that shows that, if certain seemingly plausible conditions on justice and on equality are satisfied, then the justice relation must hold that a distribution is more just if it is maximin better (i.e., makes the worst off person better off). This is a striking result, since it seems to show that a plausible egalitarian theory must overlap very significantly with the prioritarian theory of leximin.

The result rests on one further condition concerning the measurement of equality. It concerns the effect on equality of increasing the benefits of all the worst off persons while decreasing the benefits of all the best off persons. If there is only one worst off and only one best

off person (i.e., if there are no ties for worst off or for best off), then the condition concerns changes in only those two people. If, however, there is a tie for worst off, or for best off, then the condition concerns the same change for *all* the worst off, and *all* the best, off. Consider then:

Strong Conditional Contracting Extremes: Equality is increased by the combined effect of (1) increasing the benefits of all the worst off individuals without causing any of them to cease to be a worst off person, and (2) decreasing the benefits of all the best off individuals without causing any of them to cease to be a best off person.

The rough idea is that equality is increased by contracting the extremes of a distribution (reducing the benefits for those at the very top, while still leaving them at the top, and increasing the benefits at the very bottom, while still leaving them at the bottom, without any changes to anyone else). This condition judges $\langle 2, 2, 5, 8, 8, 8 \rangle$ to be more equal than $\langle 1, 1, 5, 9, 9, 9 \rangle$.

This condition is highly plausible. Indeed, it seems to be part of the concept of equality—and not merely part of plausible, but debatable, conception of equality for the theory of justice.

Bertil Tungodden has established the following:

Observation 2: If (1) the justice relation satisfies Ordering (more minimally: transitivity and reflexivity), Strong Pareto, Anonymity, and Weak Egalitarianism, and (2) the equality relation satisfies Strong Conditional Contracting Extremes, then x is more just than y if it is *maximin* better (i.e., makes the worst off person better off).³²

This result doesn't establish that leximin is the only plausible form of Paretian egalitarianism, but it yields something very close. If the conditions of the theorem are satisfied,

then (as with leximin) making the worst off person better off always makes things more just. The result leaves open how two distributions are ranked by justice in the case where the worst off person is equally well off in each distribution (whereas leximin then requires the ranking to be based on how well off the second worst off person is, etc.).

The main problem with leximin—it's near monomaniacal concern for the worst off—arises with its maximin core as well. Hence, this maximin result is troubling for egalitarians who want to avoid this feature. It appears that Weak Egalitarianism (which holds that equality prevails between anonymously Pareto incomparable distributions) cannot plausibly be combined with Strong Pareto (a weak efficiency condition) without (given the other seemingly plausible assumptions) being committed to maximin betterness being sufficient for greater justice.

I cannot here examine this troubling result. I shall simply state my (very tentative!) belief that some of the underlying conditions are not as plausible as they seem. In particular, Anonymity for justice (rather surprisingly) may not be plausible (although it is a plausible condition for equality). In the presence of the plausible Strong Pareto, which gives priority to a weak kind of efficiency, Anonymity radically increases the kind of efficiency that is given priority, and thus radically decreases the role for equality. The role of equality is limited—not merely to cases where one distribution is Pareto incomparable to another, but more generally—to cases where one distribution is *anonymously* Pareto incomparable to the other. For example, Strong Pareto entails that $\langle 1,3,9 \rangle$ is more just than $\langle 1,3,4 \rangle$, and then Anonymity adds that it is also more just than $\langle 4,1,3 \rangle$. This latter distribution is Pareto-incomparable with $\langle 1,3,4 \rangle$ and more equal. Some egalitarians may want to hold that it is more just (or at least not less just). Anonymity in this context, however, rules out this judgement. Hence, this strong form of anonymity may be rejected for some weaker form that leaves more room for egalitarian considerations.

The rejection of Anonymity, however, may not be sufficient to avoid this troubling result. A second way of challenging the result is to challenge the framework in which it is formulated. It assumes that justice can be articulated in terms of a transitive binary relationship (of being-at-least-as-just-as), but this is not obviously so. A weaker conception of justice would only require that justice determine, for any given feasible set of distributions, which distributions are just. This approach does not require a general transitive ranking of distributions in terms of justice; it only requires the selection of just alternatives for a given feasible set. Of course, it may be that this second approach, once various plausible assumptions are made, will also produce the maximin result. Further investigation is needed.³³

In summary, until a clear case can be made for rejecting one or more of the crucial conditions in Tungodden's result (e.g., Anonymity and transitivity), it appears that it may not be possible to have a Paretian egalitarian theory of justice that is radically different from the leximin prioritarian theory of justice. This is an important issue that needs to be resolved.³⁴

6. Conclusion

Three of the main theories of justice are sufficientarianism, prioritarianism, and egalitarianism. Sen's work in this area has vastly increased our understanding of the key issues by identifying key conditions that underlie each approach. In presenting a survey of these issues, I have given indications of which approaches seem plausible. Obviously, this is all very controversial, and my brief assessment should be understood solely as the beginning of a more careful examination.

References

- Charles Blackorby, Walter Bossert, and David Donaldson, "Utilitarianism and the theory of justice." in *Handbook of Social Choice and Welfare*, Vol. 1, edited by Kenneth Arrow, Amartya Sen and Kotaro Suzumura (North-Holland, Amsterdam, 2002), pp. 543-596.
- John Broome, "Equality vs. Priority: A Useful Distinction," in Daniel Wikler and Christopher J.L. Murray, eds., *'Goodness' and 'Fairness': Ethical Issues in Health Resource Allocation* (unpublished 2003).
- Partha Dasgupta, Amartya Sen, and David Starrett, "Notes on the Measurement of Inequality," *Journal of Economic Theory* 6 (1973): 180-187.
- Claude d'Aspremont and Louis Gevers, "Equity and the Informational Basis of Collective Choice," *Review of Economic Studies* 44 (1977): 199-210.
- Robert Deschamps and Louis Gevers, "Leximin and Totalism: A Joint Characterization," *Journal of Economic Theory* 17 (1978): 143-63.
- Marc Fleurbaey, "Equality versus Priority: How Relevant is the Distinction?," in Daniel Wikler and Christopher J.L. Murray, eds., *'Goodness' and 'Fairness': Ethical Issues in Health Resource Allocation* (unpublished 2003).
- James Foster, "On Economic Poverty: A Survey of Aggregate Measures," in *Advances in Econometrics*, Vol. 3, edited by R.L. Basman and George F. Rhodes, Jr. (JAI Press, Greenwich, CT, 1984).
- James Foster, Joel Greer, and Erik Thorbecke, "A Class of Decomposable Poverty Measures," *Econometrica* 52 (1984): 761-66.
- Harry Frankfurt, "Equality as a Moral Idea," *Ethics* 98 (1987): 21-43
- Louis Gevers, "On Interpersonal Comparability and Social Welfare Orderings," *Econometrica* 47 (1979): 75-89.

- Peter J. Hammond, "Equity, Arrow's Conditions, and Rawls' Difference Principle," *Econometrica* 44 (1976): 793-804.
- Karsten Klint Jensen, "What is the Difference between (Moderate) Egalitarianism and Prioritarianism?" *Economics and Philosophy* 19 (2003): 89-109.
- Dennis McKerlie, "Egalitarianism," *Dialogue* 23 (1984): 223-238.
- Derek Parfit, "Equality or Priority?" (1991 Lindley lecture) in Matthew Clayton and Andrew Williams, *The Ideal of Equality* (New York: St. Martin's Press, 2000), pp. 81-125.
- Stuart Rachels, "Counterexamples to the Transitivity of *Better Than*," *Australasian Journal of Philosophy* 76 (1998): 71-83.
- Kevin Roberts, "Possibility Theorems with Interpersonally Comparable Welfare Levels," *Review of Economic Studies* 47 (1980): 409-420.
- Amartya Sen, *Collective Choice and Social Welfare* (San Francisco: Holden-Day, 1970).
- Amartya Sen, "Rawls versus Bentham: An Axiomatic Examination of the Pure Distribution Problem," *Theory and Decision* 4 (1974): 301-310. Reprinted with slight revisions in Norman Daniels, *Reading Rawls* (New York, Basic Books Inc., 1975), 283-292
- Amartya Sen, "Poverty: An Ordinal Approach to Measurement," *Econometrica* 44 (1976a): 219-231. Reprinted in Amartya Sen, *Choice, Welfare, and Measurement* (Oxford: Basil Blackwell, 1982), pp. 373-87.
- Amartya Sen, "Welfare Inequalities and Rawlsian Axiomatics," *Theory and Decision* 7 (1976b): 243-262. Reprinted in Robert Butts and Jaako Hintikka, *Logic, Methodology, and Philosophy of Science* (Dordrecht: D. Reidel, 1977).
- Amartya Sen, "On Weights and Measurements: Informational Constraints in Social Welfare Analysis," *Econometrica* 45 (1977a): 1539-72. Reprinted in Amartya Sen, *Choice, Welfare, and Measurement* (Oxford: Basil Blackwell, 1982), pp. 226-263.

- Amartya Sen, "Social Choice Theory: A Re-Examination," *Econometrica* 45 (1977b): 53-89.
Reprinted in Amartya Sen, *Choice, Welfare, and Measurement* (Oxford: Basil Blackwell, 1982), pp. 158-200.
- Amartya Sen, "Ethical Measurement of Inequality: Some Difficulties," in W. Kreller and A.A. Shorrocks, eds., *Personal Income Distribution* (Amsterdam: North-Holland, 1978), pp. 81-94. Reprinted in Amartya Sen, *Choice, Welfare, and Measurement* (Oxford: Basil Blackwell, 1982), pp. 416-431.
- Amartya Sen, "Issues in the Measurement of Poverty," *Scandinavian Journal of Economics* 81 (1979): 285-307.
- Amartya Sen, *Commodities and Capabilities* (Amsterdam: North Holland, 1985).
- Amartya Sen, "Social Choice Theory," in *Handbook of Mathematical Economics*, ed. Kenneth Arrow, and M.D. Intriligator, vol. III (Amsterdam: North Holland, 1986), 1073-181.
- Amartya Sen, *On Ethics and Economics* (Oxford: Basil Blackwell Ltd., 1987).
- Amartya Sen, *Inequality Reexamined* (Cambridge, MA: Harvard University Press, 1992).
- Amartya Sen, *On Economic Inequality*, Enlarged Edition (Oxford: Oxford University Press, 1997).
Original edition published in 1973.
- Amartya, Sen, *Rationality and Freedom* (Cambridge, MA: Harvard University Press, 2002).
- Larry Temkin, "Intransitivity, and the Mere Addition Paradox," *Philosophy and Public Affairs* 16 (1987): 138-87.
- Larry Temkin, "A Continuum Argument for Intransitivity," *Philosophy and Public Affairs* 25 (1996): 175-210.
- Larry Temkin, *Inequality* (Oxford: Oxford University Press, 1993).
- Bertil Tungodden, "Egalitarianism: Is Leximin the Only Option?," *Economics and Philosophy* 16 (2000): 229-246.

Bertil Tungodden, "The Value of Equality," *Economics and Philosophy* 19 (2003): 1-44.

Bertil Tungodden and Peter Vallentyne, "On the Possibility of Paretian Egalitarianism," *Journal of Philosophy* 102 (2005): 126-54.

Peter Vallentyne, "Equality, Efficiency, and Priority for the Worse Off", *Economics and Philosophy* 16 (2000): 1-19.

Notes

¹ Sen's general toleration of indeterminacy reflects an extremely important point. For any given subject matter (and especially when it is normative), many questions may not have completely determinate answers. Furthermore, the mere presence of some indeterminacy does not entail radical indeterminacy. For discussions of this point, see Sen (1985, p. 205; 1992, p. 47, 134).

² See, for example, Sen (1992), pp. 7-8, 87, 92, 136, and 145-6.

³ See, for example, Temkin (1987), Temkin (1996), and Rachels (1998).

⁴ The need to reject the standard strong form of Anonymity is implicit in Sen's important result on the impossibility of the Paretian liberal. The liberalism condition of this result requires that each person have privileged jurisdiction over certain matters (e.g., when one goes to sleep), which violates Anonymity. See, chs. 6 and 6* of Sen (1970).

⁵ By way of contrast, Weak Pareto holds that x is more just than y if *everyone* has more benefits in x than in y . It is silent if even just one person has equal benefits. Weak Pareto is incredibly weak and thus highly plausible. Strong Pareto is stronger but still highly plausible (although of course somewhat less so).

⁶ In the context of Transitivity and Anonymity, Pigou-Dalton Transfer is equivalent to a condition known as "strict S-concavity". For discussion, see Sen (1997, pp. 51-56) and Dasgupta, Sen, and Starrett (1973).

⁷ Other kinds of distribution-sensitive principles include desert principles (favoring distributions that better ensure that people get what they deserve) and entitlement principles (favoring distributions that better ensure that people get what they have a right to get). Although Sen has discussed such principles, his main contribution to the topic of distribution-sensitivity concerns sufficientarianism, egalitarianism, and prioritarianism.

⁸ An important statement of sufficientarianism is Frankfurt (1987).

⁹ Appendix 7 of Sen (1997) and Sen (1979a) are terrific summaries of Sen's work on poverty measurement, as well as that of others. See also, Sen (1976a), Ch. 7 of Sen (1992), and Foster (1984).

¹⁰ I'm indebted to Walter Bossert for pointing this out to me.

¹¹ If the criterion of non-poverty (adequacy) is societally relative and depends on the overall benefits in society, then increasing the benefits of the non-poor may increase the poverty line and thereby affect the measure of poverty. Focus does not rule this out.

¹² Here we limit our attention to pure sufficientarian theories. Impure versions can be more demanding. For example, one could hold that sufficientarian concerns are *prior* to all others, but when two distributions are equally good in ensuring that people get adequate benefits, then other concerns (such as efficiency and a stronger distribution-sensitivity) apply. See Tungodden (2003) for a defense of this kind of approach.

¹³ Sen (1976a, p. 223) states this assumption as: When all the poor have the same income (no inequality among the poor), then the poverty ranking is based on the head count ratio multiplied by the poverty gap ratio. In the context of the first assumption (ranking based on the normalized weighted sum of the poverty gaps), this is equivalent to normalizing by dividing by population size.

¹⁴ More precisely: Leximin poverty gap is not representable as a weighted sum of poverty gap, if, as is usual, poverty gaps are measured by standard numbers (e.g., involving no infinitesimals). If non-standard numbers are used, and I believe this would not be inappropriate, then it is so representable.

¹⁵ Sen (1997, p. 176) notes that the following modification, developed by Shorrocks, of Sen's measure satisfies Strong Transfer Axiom: $S^*(y) = [HCR(y) \cdot PGR(y)] + [1 - HCR(y) \cdot PGR(y)] \cdot$

$G_p(y)$. Sen (2002, p. 89, fn. 40) states that he finds this modified measure more plausible.

¹⁶ Sen (1997) contains references to many other important works on poverty measurement. A particularly important class of measures (satisfying certain additive decomposability and subgroup consistency conditions) has been developed by Foster, Greer, and Thorbecke (1984).

¹⁷ For the classic statements of prioritarianism, see McKerlie (1994) and Parfit (1991).

¹⁸ See, for example, Sen (1974, 1976b, 1986, 1987a).

¹⁹ Fleurbaey (unpublished, 2003) has recently challenged whether there is a principled difference between prioritarianism and impure forms of egalitarianism that hold that justice is increased by Pareto improvements (even though they decrease equality). He shows that any purported prioritarian theory can be represented by the product of total benefits and a factor that measures equality. This mere representability, I would argue, is not sufficient to show that a theory is egalitarian in any interesting sense. Instead, we should reserve the term “egalitarian” for those theories that satisfy Pigou-Dalton *but violate Strong Separability*. Prioritarian theories, by contrast, are those theories that satisfy Strong Pareto, Pigou-Dalton, and Strong Separability. For excellent discussions of prioritarianism’s commitment to Strong Separability, see Tungodden (2003), Klint Jensen (2003), and Broome (unpublished, 2003).

²⁰ See Sen (1977a, p. 1549). The result is drawn from the work of Deschamps and Gevers (1978), Gevers (1979), and Roberts (1980). All the results noted here and below implicitly assume: (1) *Benefitism*: The justice ranking relation is defined in the space of benefits (so that non-benefit features, such as rights violations, are not relevant), (2) *Universal Domain*: Benefit distributions can take any logically possible shape. (3) *Independence of Irrelevant Alternatives*: The ranking of two alternatives depends only on their respective distributions of benefits (and not on the features of other distributions). Also, to keep things simple, I invoke Pigou-Dalton rather than the much weaker original condition of Minimal Equity, which requires that for at

least two distributions, x and y , in which (1) the best off person is the same in each distribution, (2) the best off person is better off in y than in x , and (3) everyone else is better off in x than in y , x is at least as just as y .

²¹ See Sen (1977a, p. 1547). The result is based on Hammond (1976) and D'Aspremont and Gevers (1977).

²² Unlike the other conditions I formulate, this one and Extended Pigou-Dalton are my own inventions and have not, as far as I know, been systematically studied.

²³ In the social choice literature, this condition is known as “Cardinal Full Comparability Invariance”.

²⁴ See Theorem 5 of Sen (1977a, p. 1549), which is based on Deschamps and Gevers (1978). I have added the conditions of Minimal Aggregative Efficiency and Pigou-Dalton.

²⁵ This result is from Blackorby, Bossert, and Donaldson (2002), and assumes that there are at least three individuals. They call strong prioritarian totalism “generalized utilitarianism”.

²⁶ The following theory satisfies all the above conditions: a distribution is at least as just as another if and only if (1) the total weighted benefits received by the poor (given specified poverty line) is greater, or (2) the total weighted benefits received by the poor is the same, and the total weighted benefits received by the non-poor is at least as great. This theory is just like strong prioritarian totalism except that it gives lexical priority to benefits to the poor over the non-poor. I'm indebted to Bertil Tungodden for this example.

²⁷ See, for example, Sen (1978, 1992, and especially 1997) and Dasgupta, Sen, and Starrett (1973). See also Temkin (1993) for an excellent philosophical discussion of the issues in equality measurement.

²⁸ See, for example, Sen (1997, p. 54) and Dasgupta, Sen and Starrett (1973). Note that for a fixed mean and population-size of n Lorenz dominance is equivalent to second order dominance:

for all integers i between 1 and n , the total benefits for the poorest i people is at least as great, and for some such i it is greater. The result is based on a result from Hardy, Littlewood, and Polya in the 1930s.

²⁹ I introduce this condition in Vallentyne (2000), in which I defend a particular conception of equality. I there defend a strengthening of this condition, which drops the requirement that the distance between the donor and recipient be the same for the upward transfer as for the downward transfer.

³⁰ It's worth noting, however, that the Gini coefficient, a common measure of inequality violates the weakened diminishing transfer condition.

³¹ A more careful statement of this condition would make it conditional on there being a natural zero for benefits and on all benefits being non-negative or all being non-positive.

³² See Tungodden (2000). He does not there explicitly appeal to Weak Egalitarianism as a condition of justice or to Strong Conditional Contracting Extremes as a condition on equality. Instead, he combines them into one condition on justice.

³³ The general issue of whether social choice (e.g., justice) should be understood as based on a binary ranking relation or on a choice function (that selects from the feasible set) has been extensively studied by economists. In particular, it's well known that if certain expansion and contraction consistency properties are accepted, then the choice function approach will reduce to the binary relation approach. See, for example, Sen (1977). In Tungodden and Vallentyne (in progress, 2003), we examine whether these properties should be accepted.

³⁴ For helpful comments and criticisms, I'm indebted to Dick Arneson, Walter Bossert, Ingrid Robeyns, Bertil Tungodden, Paul Weirich, and Andrew Williams.