

Chapter 5 Difference Principlesⁱ

Philippe van Parijs

(Published in *The Cambridge Companion to John Rawls*,
Samuel Freeman ed., Cambridge University Press, 2003, pp. 200-240.)

Few components of John Rawls's political philosophy have proven so epoch-making as what he somewhat oddly called the "difference principle".ⁱⁱ None has exercised as great an influence outside the circle of academic philosophers. And hardly any has given rise to so many misunderstandings or generated such heated controversies.

The core of the principle is a simple and appealing idea: that social and economic inequalities should be evaluated in terms of how well off they leave the worst off. The idea is simple: it amounts to asking that the minimum of some index of advantage should be maximised. To many, it is also appealing, for the demand that the advantages enjoyed by the least advantaged should be as generous as (feasibly) possible provides a transparent and elegant way of articulating an egalitarian impulse and a concern for efficiency. For it avoids, at the same time, the absurdity of equality at any price and the outrageousness of maximising the aggregate no matter how distributed.

Thus understood, the difference principle bears some undeniable resemblance to the justification of economic inequalities by reference to some notion of the general interest, as in the utilitarian tradition. But aggregate social welfare is not quite the same as the interest of the least advantaged. The idea of using the latter as the benchmark for assessing inequalities had never been given, before Rawls, a powerful explicit formulation that could capture the scholarly imagination. But it had occurred to others before him. For example, the famous latin fable of the stomach and the limbs is said to have been used by Roman patricians to justify their privileges to the populace: it is in the limbs' interest to feed the stomach through their hard work, for with an empty stomach, the limbs would soon wither.ⁱⁱⁱ In Anatole France's novel *L'Ile aux Pingouins* a rich farmer argues along the same track that "that little should be asked from those who possess a lot; for otherwise the rich would be less rich and the poor would be poorer."^{iv} Rawls's own

single reference to an earlier statement of the idea is to a contemporary passage by George Santayana, to the effect that "an aristocratic regime can only be justified by radiating benefit and by proving that were less given to those above, less would be attained by those beneath them."^v

Once given a bold and careful formulation, the difference principle did so well at arousing the interest of philosophers, economists and other social scientists that there is not the slightest prospect of this chapter doing justice to the huge secondary literature it generated.^{vi} Against the background of a succinct presentation of Rawls's main formulations, I shall limit myself to a clarification of some frequent misunderstandings and a brief discussion of some major difficulties in connection with three central issues. What is the exact content of the maximin criterion? Does the difference principle apply it to outcomes or to opportunities? How widely should one conceive the realm of possibilities over which maximization operates?

I. The criterion

A. Rawls's two conceptions of justice

The difference principle is one of the three components of Rawl's "special conception of justice." The latter is distinguished from his "general conception," which applies when, owing in particular to inadequate economic development, "social conditions do not allow the effective establishment of [basic liberties]" (TJ 152/132 rev). According to this general conception,

[a]ll social values X liberty and opportunity, income and wealth, and the bases of self-respect X are to be distributed equally unless an unequal distribution of any, or all, of these values is to everyone's advantage. Injustice, then, is simply inequalities that are not to the benefit of all. (TJ 62/54 rev.)

The "special conception" is "the form that the general conception finally assumes as social conditions improve. (TJ 83/ passage omitted p. 72 rev.). The crucial difference is that the special conception assigns strict priority to the equal distribution of basic liberties and to fair equality of opportunities over the demand that other "social values" be distributed "to everyone's advantage".

It is this special conception which is encapsulated in Rawls's famous two principles of justice, which A Theory of Justice first expresses as follows:

First: each person is to have an equal right to the most extensive basic liberty compatible with a similar liberty for others [the first principle, or a principle of equal liberty].^{vii}

Second: social and economic inequalities are to be arranged so that they are both (a) reasonably expected to be to everyone's advantage [the first part of the second principle, or the difference principle], and (b) attached to positions and offices open to all [the second part of the second principle, or the principle of fair equality of opportunity]" (TJ 60/ cf. 53 rev.) [Rawls introduces the labels later in TJ.]

Subject to the prior constraint of the other two principles, the difference principle thus requires that one should "arrange social and economic inequalities so that everyone benefits." (TJ 61/53 rev.)^{viii} Hence, the general conception of justice "is simply the difference principle applied to all primary goods including liberty and opportunity and so no longer constrained by other parts of the special conception." (TJ 83/cf. 73 rev., passage omitted) As used by Rawls and as discussed here, the difference principle applies only to a subset of social primary goods X the "social and economic advantages" X and within the confines set by the distribution of the other social primary goods X basic liberties and opportunities X stipulated by the other two principles.

B. A Theory of Justice's three formulations of the difference principle

I shall focus below on the nature of the primary goods contained in this subset, but want to concentrate first on the criterion of distribution which the difference principle proposes. What does it mean for social and economic inequalities to be arranged so as to be "to everyone's advantage"? This is an "ambiguous phrase", Rawls notes (TJ 61/53 rev.), which his more explicit "second formulation" is meant to clarify:

Assuming the framework of institutions required by equal liberty and fair equality of opportunity, the higher expectations of those better situated are just if and only if they work as part of a scheme which improves the expectations of the least advantaged

members of society. The intuitive idea is that the social order is not to establish and secure the more attractive prospects of those better off unless doing so is to the advantage of those less fortunate. (Rawls TJ 75/65 rev.)

In order to spell out the criterion of distribution which emerges from this second formulation, let us first suppose that our society comprises only two categories of people X the "more fortunate" and the "less fortunate," distinguished by the amount of social and economic advantages the features they possess enable them to expect. Let us further suppose, for the time being, what Rawls (TJ 81-2/71-2 rev.) calls "close-knitnes", i.e. that changes in the expectations of the better off always affect, negatively or positively, the expectations of the worse off. In this context, we can distinguish the following two jointly exhaustive situations of inequality:

- (1) *The less fortunate currently enjoy worse prospects than they would under equality.*
- (2) *The less fortunate currently enjoy better prospects than they would under equality.*

If the context was zero-sum, i.e. if the issue boiled down to dividing an existing cake, then only option (1) could arise, and any existing inequality between the more fortunate and the less fortunate would necessarily be unjust, according to the difference principle. But if it is not zero-sum, i.e. if the relevant cakes need to be made and remade and if their size can therefore be affected, through some mechanism, by how they are divided, then it becomes possible for the better prospects of the better off to improve those of the worst off. How? The most frequently mentioned possibility is that they "act as incentives so that the economic process is more efficient, innovation proceeds at a faster pace, and so on" (TJ 78/68 rev.). But this is not the only reason, probably not even the main reason, why maximin may diverge from equality. Inequalities can also be "a way to put resources in the hands of those who can make the best social use of them."^{ix} The most cogent efficiency-based case for capitalist inequalities (it is sometimes argued) does not rest on the fact that the expectation of huge gains lures entrepreneurs into working hard and taking risks, but on the fact that capitalist competition keeps removing wealth, and hence, economic power from those who have proved poor innovators or unwise investors, while concentrating it in the hands of those who find and keep finding the cheapest ways of producing the goods that best satisfy consumer demand.^x This mechanism would be destroyed if profits were redistributed in egalitarian fashion or collected by a public agency. Thus, inequalities

of income and wealth may be no less significant as enabling devices than as incentives. This possibility is even more obvious in the case of inequalities of powers and prerogatives attached to social positions.

Both the incentive mechanism and the enabling mechanism make room for (2), i.e. for the possibility that inegalitarian schemes may "improve the expectations of the least advantaged members of society, relative to what they would be under equality. According to Rawls's characterisation quoted above, such an improvement provides, subject to the other two principles of justice being satisfied, a necessary and sufficient condition for an inequality between the more fortunate and the less fortunate to be just. In other words, all that is needed to justify an inequality, however large, is some improvement, however tiny, for the worse off, relative to the conceivably very depressed counterfactual situation of total equality between the expectations of the more fortunate and the less fortunate. If this were the end of the story, as implied in many casual formulations, Rawls's difference principle would be extremely tolerant of social and economic inequality. But it is not. A Theory of Justice contains a third and far more demanding formulation of the difference principle, which is also the one Rawls routinely uses, with minor variants, in later statements:^{xi} "Social and economic inequalities are to be arranged so that they are both (a) to the greatest benefit of the least advantaged and (b) attached to offices and positions open to all under conditions of equality of opportunity . " (TJ 83/72 rev., my emphasis) This far more stringent condition for just inequalities forces us to distinguish, among situations of type (2) above ("The less fortunate currently enjoy better prospects than they would under equality"), the following three mutually exclusive cases:

(2a) The prospects of the less fortunate could be better than they are now if the prospects of the more fortunate were worse ("unjust").

(2b) The prospects of the less fortunate would be worse than they are now if the prospects of the more fortunate were worse, but they could be better if the prospects of the more fortunate were better ("just throughout").

(2c) The prospects of the less fortunate would be worse than they are now both if the prospects of the more fortunate were worse and if they were better ("perfectly just").

This final formulation involves a shift from the requirement of "some improvement" to the requirement of a "maximal improvement." It thereby opens up the possibility of inequalities which benefit the less fortunate (relative to full equality) but are nonetheless unjust by excess: lowering the expectations for the better off can raise the expectations of the worse off (2a). It also opens up the possibility of inequalities which benefit the less fortunate and which Rawls does not want to call unjust but which he regards nonetheless as suboptimal by default, as less just than greater inequalities would be: "even higher expectations for the more advantaged would raise the expectations of those in the lowest positions" (2b). Rawls describes schemes which satisfy the latter description as "just throughout, but not the best just arrangement." ("Distributive Justice," p.138; TJ 79/68 rev.) Still assuming close-knitnes between the expectations of the more fortunate and those of the less fortunate, it is therefore not enough, for a scheme to be just, that it should satisfy the broad condition (2). Far more restrictively, it must satisfy the disjunction of (2b) and (2c). And for it to be, in Rawls's phrase a "perfectly just scheme" (ibid.), it must satisfy the far more stringent (2c). In other words, however much a scheme improves the expectations of the less fortunate relative to full equality, it cannot be just if a more egalitarian scheme could sustainably do better for the less fortunate. It can be just, but not perfectly just, if a more inegalitarian scheme could sustainably do better for the less fortunate.

C. When close-knitnes breaks down: more and less egalitarian versions of the difference principle

What happens when the simplifying assumption of close-knitnes breaks down? The possibility then arises that the prospects of the less fortunate may remain unaffected by upward or downward changes in the prospects for the better off. "Perfectly just" schemes, understood as schemes relative to which the expectations of the less fortunate could not be improved, are then no longer necessarily of type (2c). For the expectations of the less fortunate to be unimprovable, it is no longer required that they should be worsened if the more fortunate had less or more than now. They can also remain unchanged. The following further three cases thereby become possible:

(2d) The prospects of the less fortunate would be worse than they are now if the prospects of the more fortunate were worse but they could remain the same as now (though not improve) if the prospects of the more fortunate were better.

(2e) The prospects of the less fortunate would be worse than they are now if the prospects of the more fortunate were better but they could remain the same as now (though not improve) if the prospects of the more fortunate were worse.

(2f) The prospects of the less fortunate could remain the same as now (though not improve) both if the prospects of the more fortunate were better and if they were worse.

According to Rawls's third and final formulation of the difference principle quoted above, schemes which satisfy any of these three conditions are as perfectly just as those which satisfy (2c). But room is now made for two more restrictive variants, one more and one less egalitarian than Rawls's "final formulation," each occasionally used by Rawls himself and each matching, under our simple two-class assumption, one of the two main justifications he gives for the difference principle. According to the more egalitarian variant, a necessary condition for perfect justice to obtain is that inequalities should contribute to improving the expectations of the less fortunate.^{xii} Unlike schemes which satisfy (2c) or (2d), schemes which satisfy (2e) or (2f) are less than perfectly just, since they contain an unused potential for reducing inequality without worsening the fate of the worse off.^{xiii}

According to the less egalitarian variant, it is not required for perfect justice to obtain that inequalities should contribute to improving the expectations of the less fortunate, but only that they should not worsen them. Indeed, perfect justice requires inequalities that improve the prospects of the better off if such improvement does not make the prospects of the worse off any worse. Of the four possible cases, it is now (2c) and (2e) that are perfectly just, while (2d) and (2f) are less than perfectly just, since they contain an unused potential for increasing the prospects of some without worsening the fate of the worse off. Unlike the first variant, this second variant is consistent with efficiency, understood as Pareto-optimality or the exhaustion of all possibilities of making some better off without making others worse off. The adoption of this

less egalitarian variant is therefore entailed by Rawls, whenever he presents the difference principle as consistent with efficiency.^{xiv}

One appealing argument in support of the more egalitarian variant goes as follows. If and only if (2c) or (2d) is satisfied is it possible to legitimise existing inequalities in the eyes of their very victims by telling them: "True, there are undeserved inequalities and you have been unlucky. But under any other feasible scheme you would be even worse off." This argument has been frequently used, including by Rawls himself when presenting the difference principle as a principle of reciprocity or mutual benefit, as a principle that "can be justified to everyone, and in particular to those who are least favored." When the difference principle is satisfied, Rawls writes, the least favoured individual "B can accept A's being better off since A's advantages have been gained in ways which improve B's prospects. If A were not allowed his better position, B would be even worse off than he is." (TJ 103/cf. 88 rev. paragraph amended). Phrased in this way, the argument is, however, incorrect. For it should practically always be possible to imagine a feasible scheme under which the particular people who currently form the worst off group would be made better off. If the absence of any such alternative feasible scheme was required by the difference principle, the latter would be unsatisfiable. But the difference principle is defined by Rawls in terms of anonymous positions, not in terms of proper names. For it to be satisfied, the worse off must not be at least as well off as they would be under any other feasible scheme, but at least as well off as the worse off, whoever they may be, would be under any other feasible scheme. Hence, when the egalitarian variant of the difference principle is satisfied, one cannot tell the less fortunate: "Under any other feasible scheme your prospects would be even worse." But one can still say: "Under any other feasible scheme someone's prospects would be even worse." The legitimising power of this correct version of the argument, its potential for reducing the "strains of commitment" of the less fortunate, is not quite as great. But it is not insignificant.

When the less egalitarian variant of the difference principle is satisfied, on the other hand, it is no longer necessarily possible to tell the victims of inequalities that under any other feasible scheme someone would be worse off than they are. The difference principle can now be satisfied even when less egalitarian schemes are feasible under which no one would be worse off (though some would be no better off) than the worse off now are. Under this variant of the difference principle, just inequalities are therefore no longer so safely protected against their

victims' "strains of commitment." On the other hand, this variant has the advantage of fitting far better into the other main argument for the difference principle, the one framed in terms of the original position. For behind the veil of ignorance, the parties "strive for as high an absolute score as possible. They do not wish a high or a low score for their opponents, nor do they seek to maximize or minimize the difference between their successes and those of others" (TJ 144-5/125 rev.). They may therefore give an absolute priority to the level of expected benefit they would end up with in case they were unlucky enough to be among the least fortunate. But when comparing two options which are equivalent on this score, they will unambiguously prefer the one that will give them more in case they turned out to be fortunate after all.

If applied to situations which are not close-knit, situations in which it is possible for (some) inequalities to be neither beneficial nor damaging for the worse off, Rawls's formulations clearly express two incompatible variants of the difference principle, each of them called for by one of the main arguments he puts forward in support of the difference principle. However, if the assumption of close-knitness does hold, (2d), (2e) and (2e) never obtain and both the egalitarian and the lexical variants yield the same selection: (2c). Rawls clearly believes that this assumption is a reasonable one to make and therefore warns that he will "always use the difference principle in the simpler form" (TJ 83/72 rev.), which corresponds precisely to this selection. Once we exclude the possibility of a tie at the bottom, both the anticipation of the strains of commitment and the reasoning behind the veil of ignorance support the choice of a scheme which generates a state of affairs about which it can be said: "The prospects of the less fortunate would be worse than they are now both if the prospects of the more fortunate were worse and if they were better."

D. The lexical difference principle and its radical extensions

Does it follow that the whole issue of what should be done when close-knitness does not hold is of no practical significance? It does if the principle is supposed to operate, as it does for Rawls, within a single generation rather than across generations (ADistributive Justice \equiv , pp. 145-7; TJ 284-93/251-58 rev., substantially amended) and within a single people rather than across nations.^{xv} For it is hard to imagine that any inequality of any importance between social positions in the same national society would not affect, be it slightly, one way or another, the

situation of the incumbents of the worse position. If one were to apply the difference principle trans-generationally, or even trans-nationally, this arguably would be a very different matter. For suppose those with the worst prospects belong to a generation that is no longer around, or live in a country whose situation we cannot hope to affect by our decisions. Then Rawls's "simpler form," which simply demands that inequalities should be to the greatest benefit of the least advantaged, gives precious little guidance: anything goes providing no reachable category is given a deal as bad as the (possibly abysmally bad) fate of the worst off unreachable one. To get some guidance, we need something like the two variants just discussed.

To see what this guidance could be, we must first work out the two variants for the general case of many categories. In the two-category case, a tie for the worse off required one to make the better off as badly off as possible in the more egalitarian version, as well off as possible in the less egalitarian version. The generalisation of the former is possible, and not altogether absurd, but it raises tricky problems which it would be better to avoid.^{xvi} Instead, the generalisation of the latter is straightforward: it consists in the "lexical difference principle", inspired by Amartya Sen^{xvii} and presented by Rawls (TJ 81-3/cf. 70-73 rev.) as a natural generalisation of his final formulation: the lexical maximin or leximin ranks situations in which the worst off are equally badly off by looking at the fate of the worst off but one, and so on in case of successive ties.^{xviii} According to this criterion, inequalities are fine as long as they do not hurt the worst off or, if the worst off category is unaffected, the worst off category but one, etc. Like its two-class version, this criterion is consistent with efficiency and fits easily into an original-position argument.^{xix} In an intergenerational or global context, the lexical difference principle would simply require social and economic advantages to be distributed to the greatest benefit of the worst off among those categories that can be affected by the choice under consideration between alternative schemes.

Rawls has offered various reasons for not using the difference principle across generations and across peoples and for elaborating instead a distinct principle of just savings and of a distinct duty of assistance among peoples. This is not the place to present or discuss these.. I shall only note the following paradox. Rawls's alternatives to radical extensions of the difference principle across generations and across peoples are unequivocally less egalitarian than such a radical extension would be. His conception of justice across generations requires that only

comparatively poorer generations should make net savings, from which richer generations will benefit: "contrary to the formulation of the difference principle, the worst off save for those better off." (A Distributive Justice \cong , p. 146) True, the rate of saving could be chosen in such a way that, as from the second generation, all generations, whether the initial poorer ones or the later richer ones would benefit relative to a stationary state. But this leaves out the first one, which will unambiguously be made worse off by any positive rate of saving than it would be under (trans-generationally egalitarian) zero growth. Similarly, his conception of justice between nations implies a duty of assistance of poorer nations by richer ones that falls far short of the transnational redistribution which a global difference principle would imply: "The levels of wealth and welfare among societies may vary, and presumably do so; but adjusting those levels is not the object of the duty of assistance." (LP 106).

Some may be attracted by Rawls's approach to distributive justice, while remaining unpersuaded by his alternatives to a transgenerational and global difference principle, in part precisely because of their inegalitarian implications. For them, it is essential to adopt a version of the principle which yields clear and meaningful consequences when extended along these two dimensions. As suggested above, the lexical version of the difference principle is best suited for this task. Paradoxically, it is the less egalitarian variant of the difference principle that offers the best chance of supporting the egalitarian strategy of boldly expanding its scope across both time and space.

II. The distribuendum

A. Outcome egalitarianism ?

After having focused on the distributive criterion for the difference principle let us now turn to what it is applied to: the distribuendum, which the difference principle X both in Rawls's "final formulation" and in the two variants subsequently explored X requires to be distributed in maximin fashion. Contrary to what some economists still sometimes call the "Rawlsian" criterion of distributive justice, the distribuendum of the difference principle has never been characterised in terms of utility, or welfare, or degree of preference satisfaction, but rather in terms of a subset of primary goods which Rawls calls "social and economic advantages". *Primary goods* were

initially conceived by Rawls as "things that every rational man is presumed to want" (TJ 62/54 rev., TJ 92/79 rev.). They were later redefined as "what persons need in their status as free and equal citizens, and as normal and fully cooperating members of society over a complete life". (TJ, revised edition,, p. xiii). Whether in the old or the new definition, they fall into two categories: natural and social. *Natural* primary goods include "health and vigor, intelligence and imagination." Although their possession is influenced by social institutions, they are not as directly under their control as social primary goods are (TJ 62/54 rev.). It does not follow, however, that inequalities in natural primary goods have nothing to do with social justice. "The natural distribution is neither just nor unjust; nor is it unjust that men are born into society at some particular position. These are simply natural facts. What is just and unjust is the way that institutions deal with these facts." (TJ 102/87 rev.) More precisely, social institutions can profoundly affect the consequences of natural inequalities through the distribution of those primary goods which are directly under their control: the *social primary goods*.

Social primary goods include fundamental liberties and opportunities for access to social positions. But in the special conception of justice, the distribution of these goods is governed, as we have seen, by the other two principles, to which the difference principle is subordinated. The remaining social primary goods, which are the preserve of the difference principle, Rawls calls *social and economic advantages*. These are defined by a list, whose canonical formulation consists in the following items: income and wealth, powers and prerogatives of offices and positions of responsibility, and the social bases of self-respect (see TJ 62/54 rev., TJ 93/80 rev., PL 181, etc.). For the maximin criterion, or any other maximization standard, to be applicable, each of these items must be measurable (if only roughly). Even in the case of income and wealth this is not straightforward, since some aspects of people's income or wealth (the quality of the physical environment they enjoy, for example) is not routinely translatable into monetary magnitudes. And as regards powers and prerogatives, or the bases of self-respect, far more would need to be said for measurement of any kind to get off the ground. Moreover, an overall index would need to be constructed in order to aggregate these various dimensions. Rawls is confident that this can be done, at least for the limited purposes which need to be served. All he needs the index for is to identify the worst off category in each social state achievable through the operation of some feasible institutional scheme consistent with the first two principles, and to

compare these indexes across such social states. But the strong correlation that tends to prevail between the various dimensions, Rawls argues, makes it unproblematic to identify the worst off. (TJ 97/83 rev.)

This is not enough to permit comparison across relevant social states, as some schemes may be designed, for example, so as to give a lower income but greater powers to the worst off category. A market economy with worker-owned firms may conceivably outperform conventional capitalism in terms of the powers and prerogatives associated with the worst position, while doing worse income-wise for everyone.^{xx} However, appeal to rational prudence from the standpoint of the least advantaged should enable us, Rawls believes, to weight appropriately the various goods to the extent necessary for the sake of applying maximin to the aggregate index (TJ 94). Contrary to what Rawls (TJ 91-92) suggests, more is required than intra-personal comparison, since the worst off need not be the same people in all relevant social states. But the interpersonal comparison involved is arguably unproblematic enough for us to be able to determine in which relevant social state the worst off enjoy the best combination of income and wealth, powers and prerogatives and the social bases of self-respect.

Let us therefore suppose that these measurement problems are manageable. What we end up with, it seems, is the requirement that we should maximize a synthetic index of the social and economic advantages enjoyed by the worst off member of the society concerned. This is a slightly clumsy variant of the economists' "Rawlsian" criterion of maximin welfare, which clearly qualifies as a form of outcome egalitarianism. This involves, however, a deep misunderstanding of Rawls's difference principle. It fails to accommodate, in particular, his repeated emphasis on stating the principle in terms of lifetime expectations of categories of people, rather than in terms of particular individuals= situations at particular times. (see TJ 64/56 rev., 71/62 rev., 73/63 rev., 98/cf.83-84 rev. passage amended, 99/85 rev, etc.) This is not simply, he insists, a convenient way of averaging scores "in order to make this criterion work" (TJ 98/cf. 84 rev. passage amended). It is a constituent feature of an otherwise thoroughly individualistic approach to justice.

B. The least fortunate and the incumbents of the worst position

How are these classes or categories to be understood? Rawls makes various suggestions.

According to one of them, the least advantaged are simply those with least social and economic advantages, the class of people with the most modest "place in the distribution of income and wealth + (TJ 96/82 rev.). The characterisation of the worst off category is then "solely in terms of relative income and wealth with no reference to social positions. Thus all persons with less than half of the median income and wealth may be taken as the least advantaged segment." (TJ 98/cf.84 rev.) Rawls does not reject this characterisation. On the contrary, he says he supposes that it "will serve well enough.≡ (TJ 98/cf. 84 rev. passage amended) Were it to be adopted, however, it would be hard to see what was gained, and easy to see what was lost, relative to a more fine-grained classification of individuals, eventually into singletons. For example, the category of people with an income less than half the median may have a higher expected income under one scheme, but with a much wider spread of incomes than under another, in such a way that the poorer half of this category has a far higher expected income under the latter the scheme than under the former. Why should the former scheme nonetheless be preferred? Settling for a rougher classification can hardly emerge as anything but a handy simplification.

Fortunately, Rawls mentions a second option, which "is to choose a particular social position, say that of the unskilled worker, and then to count as the least advantaged all those with the average income and wealth of this group, or less. The expectation of the lowest representative man is defined as the average taken over this whole class." (TJ 98/cf. 84 rev. passage amended) This too, Rawls says, "will serve well enough," even though "certainly not all social positions are relevant. For not only are there farmers, say, but dairy farmers [...]. We cannot have a coherent and manageable theory if we must take such a multiplicity of positions into account." (TJ 96/82 rev.). Do we not face again an irreducible arbitrariness, which gives no more than a pragmatic justification to stopping short of individuals? Moreover, whether the classification is arbitrary or not, is this not some clumsy form of outcome egalitarianism, an attempt to equalise (within the limits of the maximin) the income and wealth associated with social positions which people have come to occupy, at least in part, by virtue of their preferences and efforts?

No, it is not. Properly understood, the difference principle is an opportunity-egalitarian principle, and its being phrased in terms of expectations associated with social positions, rather than directly in terms of primary goods, is of crucial importance in this respect. How? The basic structure of a society, Rawls writes,

favors some starting places over others in the division of the benefits of social cooperation. It is these inequalities which the two principles are to regulate. Once these principles are satisfied, other inequalities are allowed to arise from men=s voluntary actions in accordance with the principle of free association. Thus the relevant social positions are, so to speak, the starting places properly generalized and aggregated. By choosing these positions to specify the general point of view one follows the idea that the two principles attempt to mitigate the arbitrariness of natural contingency and social fortune.≡ (TJ 96/82 rev.)

More explicitly:

"The least advantaged are defined very roughly, as the overlap between those who are least favoured by each of the three main kinds of contingencies. Thus this group includes persons whose family and class origins are more disadvantaged than others, whose natural endowments have permitted them to fare less well, and whose fortune and luck have been relatively less favourable, all within the normal range (as noted below) and with the relevant measures based on social primary goods." (Collected Papers, 258-9)

Restriction to the normal range derives from the assumption that "the first problem of justice concerns the relations among those who in the normal course of things are full and active participants in society and directly or indirectly associated together over the whole course of their life". Hence, one should, be it provisionally, abstract from the mentally defective and other "people distant from us whose fates arouses pity and anxiety" (id., 259).

Even among people in the "normal range" and even when fair equality of opportunity is realised as much as it can be consistent with the existence of the family, people differ considerably in the natural and social fortune they enjoy. What Rawls calls "liberal equality" is only bothered by inequality rooted in social circumstances. But

there is no more reason to permit the distribution of income and wealth to be settled by the distribution of natural assets than by historical and social fortune. [Ψ] For once we are troubled by the influence of either social contingencies or natural chance on the determination of distributive shares, we are bound, on reflection, to be bothered by the influence of the other." (TJ 74-75/64-65 rev.)

It does not follow, however, that one should adopt the "principle of redress", i.e. "try to even out handicaps", to eliminate differences in natural capacities. "There is another way to deal with them. The basic structure can be arranged so that these contingencies work for the good of the least fortunate." (TJ 101-2/87 rev.) How can this be achieved? Precisely by defining a social position that is fully accessible to the least fortunate in this sense and designing the institutions in such a way that the lifetime prospects, in terms of social and economic advantages, of the representative incumbent of this position are as good as they can feasibly be.

What is essential is not that all the least fortunate in this sense should occupy this social position, but that they should all have access to it. This can be secured in part through the ban on all discrimination implied by the prior principle of fair equality of opportunity, in part through making the skill requirements so low that anyone "in the normal range," however disadvantaged by nature and upbringing, can be certain to meet them. Occupying that position, rather than another more demanding one, has a major impact on a person's life prospects, assessed in terms of an index of income, wealth, powers and prerogatives, and the social bases of self-respect. Among individuals sharing the same social position, however, actual lifetime performance in terms of this index can vary considerably, as a result of events which combine chance and choice in varying, generally unassessable proportions: some become chronically sick, others give birth to a handicapped child, some get involved with an incredibly generous partner, others make an unexpected gain when selling their house, some keep buying on credit, others work overtime. Considerable variation in lifetime levels of income and wealth will result. The difference principle does not require us to equalize or maximin these outcomes, but only to maximize what the representative incumbent of the worst social position can expect,

i.e. the average lifetime index of income and wealth associated with a position accessible to all the least fortunate (in the normal range). Correctly understood, the difference principle is therefore far more responsibility-friendly (or ambition-sensitive) and therefore less egalitarian (in outcome terms) than it is often taken to be.

C. Guaranteed income or wage subsidies ? (1) Maximin scores

This clarification is of crucial importance for the more concrete institutional implications of the difference principle. For in addition to less problematic provisions for the sick, the elderly and the young, there are at least two importantly distinct ways in which one can think of implementing this principle for the category of people of working age. One, naturally suggested by Rawls's repeated characterization of the worst position as that occupied by "unskilled laborers" (e.g. TJ 78/67 rev. 80/69 rev.), consists in lifting the wages of the least productive jobs as much as is durably feasible, by means of wage subsidies. A paradigmatic example of such a scheme is Edmund Phelps's recent proposal of an hourly-wage subsidy for full-time workers, gradually phased out as the wage rate increases.^{xxi} The second possibility, naturally suggested by Rawls's explicit reference to "a graded income supplement (the so-called negative income tax),[≡] (TJ 275/ 243 rev.) consists in introducing a guaranteed minimum income and pitching it at the highest sustainable level.^{xxii} In either case, the very fact that the worse position is accessible to all X "offices and positions are open" and there is in this case no special skill requirement X means that any improvement in the expectations associated with that position will "trickle up" and the expectations associated with other low positions, as the rewards associated with these positions adjust in order to retain an adequate supply of suitably skilled people.^{xxiii}

Between some form of employment subsidy and some form of guaranteed income, between these two potential core components of the institutionalization of the difference principle, what should we choose? Perhaps one way of making it easier to come up with a clear-cut answer consists in rephrasing the question as follows: Which scheme will enable us to maximize the income (and hence the other social and economic advantages, assumed to be strongly correlated with income) of those with least income? There should

be little hesitation about the answer. True, subsidies channelled more narrowly through paid work are likely to depress less the total income to be distributed than would a less discriminating guaranteed minimum income. But this is most unlikely likely to offset the effect of the basic difference between the two schemes: people who do no work are entitled to the full guaranteed minimum income, whereas they are not entitled to any of the employment subsidy. As phrased so far, the difference principle therefore seems to unambiguously favour a guaranteed minimum income. This is taken for granted, for example, by Richard Musgrave in his early critique of the difference principle,^{xxiv} by a number of "Rawlsian" justifications of an unconditional basic income,^{xxv} and indeed by Rawls himself when expressing his embarrassment at the possibility that his difference principle would end up subsidising Malibu surfers.(PL 181-2, fn 9) It is precisely in order to block this implication that he proposes making a major change in his list of social and economic advantages, and hence of primary goods, by including leisure in the list.^{xxvi} Entitlement to the full subsidy would therefore be restricted to full-time workers, since part-time workers already enjoy some form of income in the form of supra-normal leisure, and its amount would decline proportionally all the way to zero as working time falls X which is exactly what an hourly-wage subsidy would achieve.

In this interpretation, the answer to the question of the institutionalization of the difference principle is unambiguous. Keep the original list of social and economic advantages and you will opt for a guaranteed minimum income at maximum sustainable level. Add leisure to it and you will opt for an employment subsidy, again at maximum sustainable level. This interpretation of the question would be correct, however, only if the difference principle were interpreted, in outcome-egalitarian fashion, as the maximization of the lowest level of income (and other social and economic advantages). But we have seen above that what the maximin criterion should be applied to does not consist in actual scores, but in the expected score associated with social positions. For the reasons spelled out above and following Rawls's explicit suggestion, it could be the position of those who, throughout their lives, perform nothing but unskilled jobs, i.e. jobs which only require those skills which everyone "within the normal range" can be assumed to possess. Depending on family circumstances and other contingencies more or less

independent of one's choice, the lifetime income of people in this position will vary considerably. But, under actual or feasible schemes, it is possible to determine the level of income and of the other social and economic advantages to be expected by people in this position accessible even to the least fortunate, i.e. the level its incumbents will achieve on average.

D. Guaranteed income or wage subsidies ? (2) Maximin expectations

When re-interpreted, as it should be, along such opportunity-egalitarian lines, the original version of the difference principle (which makes no reference to leisure) no longer generates any presumption in favour of a guaranteed minimum income. On the contrary, narrowly targeted employment subsidies are bound to be the winner. Why? One might want to argue that the very fact that paid work is more strongly rewarded, and hence that more paid work will be performed, must increase the surplus available for redistribution, and thereby raise, relative to what would be the case under a guaranteed minimum income scheme the lifetime income of the representative incumbent of the worse position. But this may not hold beyond the short term. For some forms of guaranteed minimum income may be far better than employment subsidies in terms of human capital formation. The latter may be fostered by a scheme more favourable to part-time work and career interruption, partly because it affords a greater opportunity to take time off for further training, but above all because it makes it less costly for parents to take time off in order to better look after their offspring at crucial periods of their intellectual and motivational development, and thereby helps the young grow up into balanced human beings and (partly therefore) effective economic agents. There is, however, a far stronger and more direct argument in support of a presumption for employment subsidies. With a given amount available for redistribution, the very fact that redistribution is focused on employment will induce the "representative man" in the worst social position to do far more paid work than under a work-independent guaranteed income, and it will thereby raise his expected income, which is simply the income people in this position will earn on average, not the maximum they "could" earn. Of course, income is not the only item on the list. But wealth is bound to vary in the same direction,

probably even in a more pronounced way. And the massive advantage thus gained by employment subsidies looks unlikely to be reversed by any disadvantage one might hope of establishing as regards the other dimensions: powers and prerogatives or the social bases of self-respect.

Thus, it is now employment subsidies, no longer a guaranteed minimum income, which the difference principle appears to clearly favour, with the original list of social and economic advantages. But the shift from a maximin of advantages achieved to a maximin of advantages expected by people in a particular social position does not make Rawls's sensible suggestion that leisure should be incorporated, on a par with income, into the relevant index, any less appropriate. Freely chosen time off work is no less important an indicator of the quality of the lifetime prospects of the worst off than their purchasing power. But if leisure is added to the list, the strong presumption in favour of employment subsidies collapses immediately. For whereas the expected income of the representative incumbent of the worst social position can safely be expected to be higher under the most suitable form of employment subsidies, her expected leisure can equally safely be expected to be more extensive under the most suitable form of guaranteed income. It may still be the case that the aggregate index of income, leisure and wealth to be expected by the worst off category is maximized by some form of employment subsidy. But the argument must be based on the differential impact of the two sorts of schemes on overall productivity and hence on the lifetime level of the income-leisure bundles and wealth to be sustainably expected by the worst off. As indicated earlier, which way this differential impact leans is by no means self-evident.

Can this indeterminacy not be resolved by considering the other items in the list of social primary goods whose distribution is governed by the difference principle? Under guaranteed income schemes, the worst off category will hold fewer jobs and on a more part-time basis than under employment subsidies. Does it follow that there will be fewer powers and prerogatives attached to the position they occupy? By no means, because the very fact that some of their income does not derive from the jobs the workers hold endows them with a bargaining power which should boost their expected level of powers and prerogatives. What about self-respect, "perhaps the most important primary good,"

which involves both "a person's...secure conviction that his conception of the good, his plan of life, is worth carrying out," and "a confidence in one's ability, so far as it is within one's power, to fulfill one's intentions" (TJ 440/386 rev.)? The key social condition for everyone to have access to self-respect in this sense is "that there should be for each person at least one community of shared interests to which he belongs and where he finds his endeavors confirmed by his association" (TJ 442/388 rev.). Although relevant associations are certainly not confined, in Rawls's eyes, to the realm of paid work, the latter is undoubtedly a central component of many people's plan of life, and work relations provide them with the community in which their endeavours are appreciated and recognised. Guaranteed minimum schemes of a sort that lock people into a situation of unemployment are therefore at a strong disadvantage. On the other hand, not all forms of employment relations are equally good at securing self-respect, and employment subsidy schemes which conspicuously mark off the jobs for the unskilled as subsidised (and hence not worth doing if the full cost needed to be paid) would also be defective.

However, not all guaranteed income schemes share the former defect, nor do all employment subsidy schemes share the latter. Does the concern with self-respect generate any general presumption in favour of the most suitable form of either? Employment subsidies are more effective at enlisting unskilled people into jobs in general, but they need not be better than well-designed guaranteed minimum income schemes at providing them with comparatively low-paid jobs which nonetheless provide incumbents with the community of shared interests which the sense of their own worth requires. For some of these schemes, such as a negative income tax or a universal basic income, also amount to implicit employment subsidies. They thereby enable people to buy themselves into a job, while the very unconditionality of the subsidy enables potential workers to be more discriminating, and thereby systematically favour those jobs which mean more to the workers than just the pay.

To sum up: It is crucial to interpret the difference principle, as one must, so that it applies to the lifetime prospects of people in the social position with the worst lifetime

prospects, a position that can be occupied by the least fortunate people ("in the normal range"), those whose innate talents and life circumstances are such that they have access to no other position. This position can be interpreted as the position of those who occupy or could occupy only those jobs with the least demanding skill requirement. The maximization of the expectations attached to this position does not justify massive means-tested transfers targeted at the poor of the sort characteristic of the welfare state. In Rawls's property-owning democracy, such transfers should be kept marginal, perhaps only for people "outside the normal range." Free, effective and compulsory basic education should do part of the job, by reducing the extent to which differences in talents and circumstances translate into inequalities of skills. But beyond that, some systematic channelling of resources towards those who can only occupy the least-skilled jobs is unavoidable. To preserve the social bases of the self-respect of the least fortunate, their receipt of these resources should not be based on the administrative identification of them as Aneedy \cong , but rather should take the form of general subsidy schemes benefitting the less well paid. If leisure is omitted from the index of expectations, we have seen that some sort of employment subsidy scheme is pretty certain to emerge as the most effective embodiment of the difference principle. But as soon as one sensibly assumes, along with Rawls, that the leisure that people in the worst position can expect should also count positively, along with the income they can expect to earn, in the index of expectations, then the strong presumption in favour of suitable employment subsidies melts away and at least some guaranteed income schemes emerge as powerful contenders. Whether someone committed to (the most consistent version of) the difference principle should favour one or the other type of scheme then depends not only on the exact specification of the variant considered, but also on the precise weighting of the various components of the index of primary goods (with income and wealth leaning one way, leisure and powers the other way, and self-respect ambivalent), and on an empirical assessment of the magnitude of the steady-state effects to be expected from different schemes along these various dimensions under specific social and economic circumstances.^{xxvii}

III. THE TOOLS

A. The constraint of the first two principles

We have so far scrutinized the criterion which the difference principle consists in applying and the variable to which it is applied. In its most consistent interpretation, the difference principle requires us to maximize the expectations, in terms of social and economic advantages, including leisure, of the representative incumbent of the social position with the lowest such expectations. Let us now turn, briefly, to a third and no less important issue. What the difference principle requires remains underspecified if one does not characterize which instruments it has available to help maximize whatever it says one must maximize. Depending on how wide a range of tools can legitimately be used, application of the difference principle may lead to a situation which is more or less favourable to the worst off and more or less close to strict equality.

From the very beginning, Rawls is very clear: the difference principle is meant as principle for institutions, or practices, rather than for particular actions or persons. An institution is defined as

a public system of rules which defines offices and positions with their rights and duties, powers and immunities, and the like. These rules specify certain forms of action as permissible, others as forbidden; and they provide for certain penalties and defences, and so on, when violations occur. (TJ 55/47-8 rev.).

More specifically, the realization of the difference principle must concentrate on "the primary subject of justice...the basic structure of society, or more exactly, the way in which the major social institutions distribute fundamental rights and duties and determine the division of advantages from social cooperation" (TJ 7/7 rev.). These major social institutions are,

the political constitution and the principal economic and social arrangements. Thus the legal protection of freedom of thought and liberty of conscience, competitive markets, private property in the means of production, and the monogamous family are examples of major social institutions.(ibid.)

Of course, the attempt to realise the difference principle as well as possible does not allow us to shape the basic structure any way we wish, but only within the constraints of the other two principles of justice. For example, freedom of thought and the universal right to vote affirmed by the first principle rule out both brainwashing as an instrument for boosting citizens' readiness to work, as well as disfranchising the rich as a way of making generous redistribution more secure. Even if institutions so shaped could reliably boost the expected level of social and economic advantage attached to the least advantaged position, the priority of the first principle would nonetheless ban them as unjust. Similarly, even if forced labour or a statutory division of roles between men and women could enhance a society's economic performance in such a way that higher expectations could reliably be secured to those confined to the worst off position, the subordination of the difference principle to the second part of the second principle X fair equality of opportunity, understood as entailing free choice of occupation (TJ 275/243 rev.) X would prohibit the discrimination this would involve.

On the other hand, one should not exaggerate the extent to which the operation of the difference principle is constrained in this direct way. For example, "the right to hold (personal) property" (TJ 61/53 rev.) constitutes only a very weak constraint on the rules that must govern the ownership and transfer of material goods. It is perfectly consistent with tax institutions that would spread wealth widely across society. Indeed, it is even consistent with "liberal socialism" (TJ 280/ 248 rev.), a property regime in which all means of production are publicly owned. It would only exclude a radical form of communism, in which all consumption goods too would be held in common. Similarly, the fundamental liberties do not give individuals an entitlement to the whole of the income which their personal talents enable them to earn on a free market. True, Rawls concedes that, should the question of the ownership of our native endowments arise, "it is persons themselves who own their endowments: the psychological and physical integrity of persons is already guaranteed by the basic rights and liberties that fall under the first principle of justice."^{xxviii} Moreover, the difference principle is presented as a principle for the fair distribution of the benefits and burdens of social cooperation, which suggests

that whatever people can produce solely by their talents, without relying on social cooperation, should be theirs to keep. But this should not prompt elaborate attempts to work out the residual that is up for redistribution. As regards the taxation of income or external wealth, this constraint is never binding. So much of what anyone can produce is conceivable only due to social cooperation (without which there would not even be language, for example) that the non-cooperative output of any individual (so far as that idea makes any sense at all) must shrink into insignificance. Maximin considerations will therefore put a ceiling on the rate at which the income of the more talented is taxed long before the point is reached at which any individual could claim with any plausibility that she could have generated all the remaining net income on her own. Consequently, one is entitled to "regard the distribution of talents as a common asset" (TJ 101/87 rev.) or even to "view the greater abilities as a social asset to be used for the common advantage" (TJ 107/92 rev.).^{xxix} Hence, it seems that there should be plenty of room for institutions to be shaped so as to increase the expectations of the worse off without violating either the fundamental liberties or fair equality of opportunity, as understood by Rawls.

B. No resources left for the difference principle ?

Nonetheless, taken literally, the priority of these two other principles would leave no resources for the difference principle to use. For what does the priority of the first principle mean? "This ordering means that a departure from the institutions of equal liberty required by the first principle cannot be justified by, or compensated for, by greater social and economic advantages." (TJ 61/cf. 53-4 rev.). Now, the "institutions of equal liberty" cannot reduce to the formal constitutional recognition of a number of fundamental rights. They must encompass those institutions which are required to make these rights effective, by protecting them against violation not only by government officials but also by private citizens. The prevention of private violations of physical integrity, for example, is far from fully achieved in any country. Preventing such acts entirely is presumably only possible at the cost of infringing some other fundamental rights, and could therefore be blocked by the first principle itself. But within the limits this imposes, far more could no doubt be done to increase the citizens' physical safety,

say by allocating more public resources to brighter street lighting, tighter police supervision or more sophisticated video cameras. As long as such a potential exists, devoting any amount of public resources to the satisfaction of the difference principle, rather than to a more effective protection of people's physical integrity, therefore implies an opportunity cost in terms of fundamental liberties for all, and hence a "departure from the institutions of equal liberty."^{xxx}

The same holds, *mutatis mutandis*, as regards the priority of the principle of fair equality of opportunity, again as soon as the latter does not reduce to requiring a formal ban on discrimination, but demands the effective neutralization of the impact of gender, race or social background on access to social positions. For example, on the background of diverse family backgrounds which give equally talented people greatly different life choices, it would always be possible to pour more money into schools, pre-school activities, after-school supervision, free extra-curricular activities, etc., in order to further reduce the gap between unequally favoured children. Again, a limit may conceivably be imposed by the best interpretation of the principle of fair equality of opportunity itself if making chances more equal ends up making them worse for all, through a major shrinking of what they are chances of getting access to. But as long as this limit is not reached, diverting any amount of money into improving the situation of the worst off in any other way would therefore constitute an unequivocal "departure from the institutions of equal opportunity."

This seems to generate an uncomfortable dilemma. Either we stick to the two priority rules, and the conditions under which the difference principle can legitimately guide us are most unlikely to be ever satisfied. Or we give up the priority rules, and the difference principle, as one of the three components of Rawls's special conception of justice, vanishes as such and is absorbed into the general conception of justice, which lumps all social primary goods together. Rawls recognises that "the distinction drawn [between three categories of social primary goods] and the ordering proposed [by the two priority rules] are bound to be at best approximations", but he believes that "under many circumstances anyway, the two principles in serial order may serve well enough." (TJ 63/55 rev.) We have just seen that in any realistic context, they do not serve well enough

if the priority rules are taken literally, as implementation of the two principles in serial order would leave the difference principle with nothing to play with and hence clash with what Rawls himself would regard as a "reasonable conception of justice".

The only way out of the dilemma involves adopting a milder, but still vigorous, interpretation of what "priority" means: to justify even a small loss as regards a prior principle, one must be able to argue that it is a necessary condition for a major gain as regards a subordinate principle. Resources which would only make a negligible difference in terms of fundamental liberties or equal opportunities can therefore be legitimately devoted to improving the expectations of the least advantaged. Of course, for this milder priority to make sense, it requires, unlike the stricter one, some general metric which applies, be it roughly, across the three categories of social primary goods. But those willing to accept the feasibility of an index of social and economic advantages X without which the difference principle cannot get off the ground X are unlikely to rule out the possibility of such a metric. And there is in any case no other way of escaping the dilemma stated above.^{xxx}

C. The thick conception of the well-ordered society

The difference principle is thus hopefully protected against the first two principles being too greedy. But what about the citizens being too greedy? The difference principle, let us recall, applies to institutions, more specifically the basic structure of society, not to the conduct of individual persons or particular organisations. So, suppose the legal framework mandated by the difference principle is in place. Might citizens driven by their private aims not try to hide as much as possible of their taxable income and wealth? Might the organizations representing the better paid workers not use their bargaining power to ward off any attempt to increase taxation at their expense? Above all, in our globalized economy, might the owners of financial and human capital not credibly threaten to move away to countries in which the net tax return on their assets is higher? And if this is the case, would the amount actually available for redistributive tasks not reduce to precious little, much of it moreover squandered on control and litigation costs?

Rawls's initial response lies in the full ideal of a society "well-ordered" by his principles of justice. Such a society is one in which everyone accepts and knows that the others accept the same principles of justice, and basic institutions satisfy and are known to satisfy these principles" (TJ, 453-4/397-8 rev.). Its citizens "have a strong and normally effective desire to act as the principles of justice require" (TJ 454/398 rev.). More explicitly: "To accept or to honour the principles agreed to means to apply them willingly as the public conception of justice, and to affirm their implications in our thought and conduct."^{xxxii} Thus, in addition to the three principles that characterise just institutions, certain principles for individuals "are an essential part of any theory of justice" (TJ 108/93 rev.).^{xxxiii} The exact content of these principles can and should be determined, as in the case of principles for the basic structure, by adopting the standpoint of the original position. Behind the veil of ignorance and against the background of the principles selected for the basic structure, including the difference principle, our concern to maximise our index of social and economic advantages in case we ended up in the worst position will presumably direct us to make the following sort of commitment: to willingly disclose all our income to the tax authorities, to refrain from leaving the country in search of a more favourable tax treatment, perhaps even (at least if leisure is not included in the index) to work as hard as possible at the most productive job we could do. Clearly, if the conduct of individuals conformed to these principles, application of the difference principle to the basic structure would lead to a situation with far better prospects for the worst off and far less unequal expectations than would be the case if individuals were guided by nothing but their self-interest. It has been argued, in particular by G.A. Cohen, that these would indeed be the implications of a consistent interpretation of the commitment to the difference principle characteristic of a well-ordered society.^{xxxiv} Unlike the common law interpretation of the difference principle, with its exclusive focus on institutions, this strict interpretation rules out as unjust any inequality deriving from incentive payments, as appropriately committed economic agents will not need to be bribed into doing what is best for the prospects of the worst off.

D. The institutional division of labour

However, this is not at all the path Rawls wants us to follow. Indeed, on his own account (PL, p. xvi), the most fundamental change in his position since TJ precisely consists in the abandonment of a conception of the well-ordered society as the conjunction of "justice as fairness" applying to institutions and "rightness as fairness" applying to individuals. The problem is that in TJ's account, "the members of any well-ordered society [...] accept not only the same conception of justice but also the same comprehensive doctrine of which that conception is a part, or from which it can be derived." But "given the free institutions that conception itself enjoins, we can no longer assume that citizens generally, even if they accept justice as fairness as a political conception, also accept the particular comprehensive view to which it might seem in TJ to belong."^{xxxv} Whereas a comprehensive moral doctrine is "one that applies to all subjects and covers all values", a political conception "focuses on the political (in the form of the basic structure), which is but a part of the domain of the moral." (ibid. 35) We are thus back to a narrow focus on the basic structure. How does Rawls justify it?

The reason initially given by Rawls is that the effects of the basic structure on people's lifelong prospects "are so profound and present from the start" (TJ 7/7 rev.), that it exerts a "profound and pervasive influence on the persons who live under its institutions."^{xxxvi} This makes sense as a criterion for the inclusion of an institution into the basic structure, but a profound impact is by no means confined to what could readily be described as an institution. The dispositions that govern people's behavioural responses to redistributive schemes would qualify just as easily as many components of the basic structure. Another reason, presented as decisive by Barry, is that the hierarchy of Rawls's principles of justice assigns priority to the freedom of occupational choice over the difference principle.^{xxxvii} But the question is no longer, as it was in the preceding section, whether in a just society the institutions mandated by the difference principle could restrict occupational freedom, but whether the citizens of a just society should use this freedom to the greatest benefit of the worst off. A third possibility is prompted by Rawls's insistence that "in designing and reforming social arrangements one must, of course, examine the schemes and tactics it allows and the forms of behavior which it

tends to encourage" X an approach he associates with Adam Smith's invisible hand and Jeremy Bentham's "artificial identification of interests" (TJ 57/49 rev.). Just as the combination of self-interested rational behavior and a battery of lump-sum taxes and subsidies is supposed to enable economists to reach the first-best maximum of their social welfare function, is not also the collaboration of a well-designed basic structure and selfish economic agents an indispensable tool for Rawls' difference principle? Alas not. Lump-sum taxation of the better endowed is no part of a Rawlsian's legitimate tool kit,^{xxxviii} and economic behaviour that is less selfish can therefore improve performance in terms of the difference principle.

The most promising justification for the focus on institutions rather lies in the second reason given by Rawls himself for making the basic structure the primary subject of justice:

Since a public conception of justice needs clear, simple and intelligible rules, we rely on an institutional division of labor between principles required to preserve background justice and principles that apply directly to particular transactions between individuals and associations. Once this division of labor is set up, individuals and associations are then left free to advance their (permissible) ends within the framework of the basic structure, secure in the knowledge that elsewhere in the social system the adjustments necessary to preserve background justice are being made.^{xxxix}

This institutional division of labour is undoubtedly convenient, as it "allows us to abstract from the enormous complexities of the innumerable transactions of daily life and frees us from having to keep track of the changing relative positions of particular individuals."^{xl} But does this convenience not come at the heavy cost of letting individual opportunistic behaviour undo what the difference principle attempts to do? In an attempt to avoid the slippery slope leading to a comprehensive moral conception, is one not back to the ever greater risk, in an increasingly secularized and globalized society, of being left with exceedingly feeble redistributive mechanisms as a result of self-seeking individual

conduct by economic agents?

E. Motivation-conscious institutional engineering

Here is what I believe to be the most promising line for a Rawlsian to take in order to avoid this further dilemma. When choosing the rules that will govern our collective life, it is important that we should choose rules that are not only simple and intelligible but also formulated in terms of facts that are objectively verifiable, so as to avoid arbitrary sanctions. This imposes limits on possible collective rules that are far narrower than those applying to individual maxims. It is precisely because of this difference that there is necessarily a gap between the level of inequality justified by Rawls's "lax" version of the difference principle and the one by Cohen's "strict" version. However, this gap is greatly reduced if one takes full account of the fact that the motivation guiding individual conduct cannot be assumed to be given exogenously, independently of institutions.

To start with, if the rules can be recognised to be fair by all because their choice was guided by the ideal of impartiality embodied in the original position, and if their implementation can be expected to be fair because of the verifiability of the conditions in terms of which they are formulated, then citizens will generally have no plausible excuse for infringing them, and one can therefore expect citizens to routinely comply with them, not out of fear of sanction, nor as a direct reflection of a commitment to a particular comprehensive moral conception, but out of allegiance to an institution they cannot help recognising as fair. This matches the thin notion of a well-ordered society, defined as a society in which "(nearly) everyone strictly complies with, and so abides by, the principles of justice,"^{xli} which survived Rawls's change of mind between TJ and PL.

Moreover, there is a wide range of other ways in which the choice of institutions can affect people's motivation and behaviour in daily life, and the consequences of this influence for the lifelong prospects of the worst off need to be factored in when assessing whether a particular combination of institutions constitutes a just basic structure. For example, legal rules regarding urban planning, health care provision or trade union organisation may foster, or instead counteract, segregation between age groups or income classes, in such a way that spontaneous solidarity between these categories is discouraged

or nurtured, with potentially momentous consequences for the expectations of the worst off. Similarly, social policies, labour market legislation and the regulation of credit and advertising may conceivably encourage or discourage to very different extents an ethos of work and thrift. Of course, as soon as leisure is included in the index of primary goods, an increased income for the incumbents of the worst position may be offset by the shrinking of their leisure. But promoting the work ethos of the more skilled and affluent is not similarly ambivalent, as part of their greater output, unlike their leisure, can be used to boost the expectations of the worst off. Most crucially, perhaps, as the world market tightens its grip and transnational mobility increases, a key feature of a just basic structure must reside in its ability to generate sufficient allegiance on the part of the most qualified, for countries not to have to engage, in order to retain them, in a damaging tax competition that would enhance inequality all around and arguably depress the expectations of the worst off in every country below what it would have been without the threat of mobility.

Thus, there is a Rawlsian alternative to both perfectionism and vacuity, to a difference principle incorporated in a comprehensive moral doctrine and one disempowered by individual opportunism. It can remain exclusively focused on institutions, but must bear in mind that institutions can be a powerful influence on individual motivation. Without overstepping the constraints imposed by the first two principles, it must not shy away from resolutely designing institutions that foster an ethos of solidarity, of work, indeed of patriotism, not of course because of the intrinsic goodness of a life inspired by such an ethos, but because of its crucial instrumental value in the service of boosting the lifelong prospects of the incumbents of society's worst position.

Conclusion

The aim of this chapter has not been to propose a comprehensive assessment of the difference principle, but only to spell out its core content, to review some of the main objections it has raised and to unfold along the way a number of distinct versions of the difference principle, more or less egalitarian, more or less work-oriented, more or less

restricted in scope. Let us briefly rehearse the main points.

1. The criterion. As characterised by the difference principle, just institutions certainly require far more than the worst off being better off than they would be in the absence of any inequality. The inequalities allowed by just institutions must make the worst off as well off as realistically possible. But do just institutions rule out any inequality that does not improve the position of the worst off, or any inequality that worsens it? Rawls says both, but should say the latter. And if he opted for this less egalitarian version, then radically extending the difference principle across generations and across nations X an option Rawls rules out in favour of far less egalitarian principles X would become far more sensible than it would otherwise be.

2. The distribuendum. The metric in terms of which the difference principle is formulated uses an index of social and economic advantages. But what needs to be maximised is not, in outcome-egalitarian fashion, the lowest score achieved, but rather, in opportunity-egalitarian fashion, the expected score of the incumbent of a social position accessible to the least fortunate. With Rawls's initial list of social and economic advantages, the sort of distributive scheme favoured by the difference principle is a guaranteed minimum income if it is understood, as it should not be, in terms of achieved scores. It becomes an employment subsidy scheme if it is understood, as it should, in terms of expectations. If leisure is added to the list of social and economic advantages, as Rawls later sensibly proposed, then the strong presumption switches from guaranteed income to employment subsidies providing one sticks to the achieved-scores interpretation. If instead one adopts, as one should, the interpretation in terms of expectations, the choice between the two types of schemes is crucially dependent on a more precise specification of the index and on some empirical conjectures.

3. The tools. Finally, what instruments can legitimately be used to raise the expectations of the worst off? The priority of the other two principles excludes some instruments that would directly clash with fundamental liberties or the ban on discrimination. But above all, if taken strictly, it makes the resources available for the difference principle shrink into insignificance, as their opportunity cost, in terms of what

could be done to better realise the first two principles, is never zero. Only a milder interpretation of the priority can save the difference principle from practical irrelevance. A second central issue is whether only institutions should be guided by the difference principle or also individual conduct. Rawls's original characterisation of the well-ordered society is arguably conducive to the latter, tighter interpretation. But his more recent writings emphasise the institutional focus. Just institutions will then countenance more inequalities than would otherwise be the case, but not quite as many as would be the case if individual dispositions were taken as given.

Thus, the difference principle comes in many variants, depending on choices made along many dimensions. Some of these choices are clearly, consistently and rightly made by Rawls. In other important dimensions, his choices are not so clear, not so consistent, and/or not so clearly right. Depending on what choices are made, one will end up with hardly any just inequality or a tremendous amount. Depending on what choices are made, one also ends up with a principle that does not make much sense if applied on a world scale or one that, instead, as globalisation slowly grinds mankind into one reasonably pluralistic people, will become the central reference for discussing global, not just increasingly obsolete national justice.

Notes: Difference Principles

- i. I am grateful to Andrew Williams, Samuel Freeman and to participants in the Hoover Chair's seminar in social and political philosophy for useful comments.
- ii. Rawls's first use of the expression is in Rawls ("Distributive Justice", Philosophy, Politics and Society (P. Laslett & W.G. Runciman eds), Oxford: Blackwell, 1967, 58-82. Reprinted in Collected Papers [henceforth CP], edited by Samuel Freeman (Cambridge (Mass): Harvard University Press, 1999), 130-53, at p. 138) to refer to the first part of the second principle "which we may, for obvious reasons, refer to as the difference principle". His motivation for the choice of the expression is more explicit in Rawls ("Distributive Justice: Some Addenda", in Collected Papers, 154-75, at p.163): "All differences in wealth and income, all social and economic inequalities, should work for the good of the least favored. For this reason I call it the difference principle."
- iii See e.g. Jacques Leclercq, *De la communauté populaire* (Paris: Cerf, 1938), p.53.
- iv (Paris: Calmann-Lévy, 1907) pp.83-4.
5. George Santayana, Reason and Society (New York: Charles Scribner, 1905) pp. 109-10, quoted in Rawls, ADistributive Justice≡, CP, p.151, and TJ 74/64 rev., fn12). In the context of his first extensive discussion of the difference principle, Rawls (ADistributive Justice≡, p.151) describes this passage by Santayana as "the nearest statement known to [him]," but also mentions Christian Bay, The Structure of Freedom (Palo Alto: Stanford University Press, 1958), at pp. 59, 374-5, "who adopts the principle of maximizing freedom, giving special attention to the freedom of the marginal, least privileged man".
- vi. Comprehensive presentations of the difference principle can be found in Brian Barry, The Liberal Theory of Justice. A critical examination of the principal doctrines in "A Theory of Justice" by John Rawls (Oxford: Oxford University Press, 1973); Tom Beauchamp ("Distributive Justice and the Difference Principle", in John Rawls's Thory of

Justice. An Introduction, edited by H. Gene Blocker & Elizabeth H. Smith (Athens: Ohio University Press, 1980), 132-161); Thomas W. Pogge, Realizing Rawls (Ithaca & London: Cornell University Press, 1989); Walter A Schaller, „ARawls, the Difference Principle, and Economic Inequality,“ Pacific Philosophical Quarterly 79 (1998): 368-91; Wilfried Hinsch, „ARawls= Differenzprinzip und seine sozialpolitischen Implikationen,“ in Sozialpolitik und Gerechtigkeit, eds. Blasche, Siegfried & Döring, Dieter, (Frankfurt, Germany: Campus, 1999), 17-74. Influential critiques include Nozick, Robert, Anarchy, State and Utopia (Oxford: Blackwell, 1974); Jeremy Waldron, „AJohn Rawls and the Social Minimum,“ Journal of Applied Philosophy 3 (1986): 21-33; Amartya Sen, Inequality Reexamined. (Oxford: Oxford University Press, 1992); G.A. Cohen, "Incentives, Inequality and Community", in The Tanner Lectures on Human Values Vol. XIII, edited by G.B. Peterson (Salt Lake City: University of Utah Press), 1992, 261-329, and „Where the Action Is: On the Site of Distributive Justice,“ Philosophy and Public Affairs vol. 26 (1997), 3-30.

vii In the revised edition of TJ the first principle is revised to read: „Each person is to have an equal right to *the most extensive scheme of equal basic liberties* compatible with a similar scheme of liberties for all.“ TJ 53 rev. (Emphases added.)

viii. This first formulation in TJ matches Rawls's first published formulation of the (then still unlabeled) difference principle in Rawls ("Justice as Fairness,“ Collected Papers, 47-72, at p.48): "inequalities are arbitrary unless it is reasonable to expect that they will work out for everyone's advantage."

ix Rawls

x See for example Jan F. Narveson, "A Puzzle about Economic Justice in Rawls's Theory,“ Social Theory and Practice 4 (1976):1-27, at p.10; N. Scott Arnold, "Capitalists

and the Ethics of Contribution", Canadian Journal of Philosophy 15 (1985): 87-102, at p.98.

xi. See, for example, Rawls, AA Kantian Conception of Equality, Collected Papers, 254-66, at p.258): "Social and economic inequalities [...] must be (a) to the greatest expected benefit of the least advantaged [...]."; or Political Liberalism [henceforth PL] (New York: Columbia University Press, 1993), pp. 6-7: "the social and economic inequalities attached to offices and positions are to be adjusted so that, whatever the level of those inequalities, whether great or small, they are to the greatest benefit of the least advantaged members of society.≡

xii Here are some typical statements: "differences in wealth and income are just only if they are to the advantage of the representative man who is worse off. " (ADistributive Justice≡: p.163); "According to the difference principle, [this kind of initial inequality in life prospects] is justifiable only if the difference in expectation is to the advantage of the representative man who is worse off [Ψ]". (TJ 78/68 rev.); "Those who have been favored by nature, whoever they are, may gain from their good fortune only on terms that improve the situation of those who have lost out." (ADistributive Justice,≡ p. 140; TJ 101/87 rev.); "The two principles are equivalent [Ψ] to an undertaking to regard the distribution of natural abilities as a collective asset so that the more fortunate are to benefit only in ways that help those who have lost out." (TJ 179/cf. 156 rev., adding Aregard...*in some respects* as a collective asset...≡)

xiii. Under the simplifying assumption of two categories (the more fortunate and the less fortunate), it is necessarily the case that making some of the better off better off still,

while the fate of the worse off remains unchanged, increases inequality. With more than two categories, however, inequality could decrease as a result if the better off who are benefited are not among the best off. Applying the criterion would then require a precise index of inequality, and benefiting the worst off category but one is then likely to be required when the worse off category cannot be made better off, as otherwise one would be tolerating an inequality which does not contribute to making the worst off better off.

xiv For example: "First, the difference principle satisfies the principle of efficiency."

(Rawls, *A Distributive Justice*, p. 163). "The problem is to [Ψ] find a conception of justice that singles out one of these efficient distributions as also just." (TJ 71/61 rev.); "But it should be noted that the difference principle is compatible with the principle of efficiency." (TJ 79/69 rev.); "This is illustrated clearly in the case where there are only two relevant classes; here maximin selects the (Pareto) efficient point closest to equality." ("Reply to Alexander and Musgrave," *Collected Papers*, 232-53, at p. 247.

xv See Rawls, "The Law of Peoples," *Collected Papers*, 529-64, at 558-9; *The Law of Peoples* [LP] (Cambridge, Mass.: Harvard University Press, 1999) pp. 115-19.

xvi. The generalisation of the more egalitarian variant states that inequalities are fine only if they contribute to improving the fate of the worse off. It does not follow from this criterion that one should level the situation of all reachable categories down to that of the lowest unreachable one. With any plausible index of inequality, improving the situation of the worst off category but one, or of other categories higher up which are still pretty badly off, is most likely to reduce inequality rather than to increase it, and the criterion amounts, in case of a tie at the bottom, to requiring that inequality should be minimised. A situation

which satisfies this criterion will generally not be Pareto-optimal, and its selection cannot be justified from the standpoint of the (envy-free) original position. Nor can it now be said to the worst off: "you can regard existing inequalities as fair because someone would be even worse off than you are now under any other feasible arrangement", but only: "you can regard existing inequalities as fair because someone would be even worse off than you are now under any other more egalitarian feasible arrangement". In an intergenerational or global context, the recommendation following from this criterion would be to level down if the only reachable categories are among the best off. But if they include some of the worst off, it may not be very different from what follows from the lexical difference principle to be considered below. Secure prescriptions, however, require a precise specification of the inequality index.

xvii See Sen, Collective Choice and Social Welfare, (Amsterdam: North-Holland, 1979), p. 138)

xviii. The term "lexical" is motivated by the way in which words are ordered in a dictionary: the rank of the first letter of each word takes strict priority over the rank of its second letter, etc., when deciding which word should come first.

xix. Derek Parfit (Equality or Priority?, The Lindley Lecture, University of Kansas: Department of Philosophy, 1995, p. 38) points out this connection between the lexical variant of the difference principle and the original-position argument. This lexical variant is the most egalitarian among the principles which fit Parfit's "priority view": "Benefiting people matters more the worse off the people are." (ibid. 19). Unlike the more egalitarian variant of the difference principle (in both its two-class and general form), it does not require any commitment to his "principle of equality": "It is in itself bad if some people are worse off than others." (ibid. 4)

xx. See e.g. Richard Krouse, & Michael McPherson, "Capitalism, 'Property-Owning Democracy', and the Welfare State", in Democracy and the Welfare State, edited by Amy

Gutmann (Princeton, NJ: Princeton University Press, 1988), 78-105.

xxi. See Edmund Phelps, Rewarding Work (Cambridge, Mass.: Harvard University Press, 1997).

xxii. Rawls was quick at picking up the idea of negative income tax. Milton Friedman briefly sketched it in his *Capitalism and Freedom* ((Chicago, University of Chicago Press, 1962) 191-4, but his first widely available article on the subject ("The Case for the Negative Income Tax", National Review, March 7, 1967, 239-41) and the first technical article by James Tobin, Joseph A Pechman. and Paul M. Mieszkowski, ("Is a Negative Income Tax Practical?", *The Yale Law Journal* 77 (1967), 1-27) were published in the same year as Rawls's first article on the difference principle, in which the description of the just social structure ends as follows: "Lastly, there is a guarantee of a social minimum which the government meets by family allowances and special payments in times of unemployment, or by a negative income tax" (ADistributive Justice, CP,p.41).

xxiii. See TJ 82/71 rev. This instantiates what Rawls calls "chain connection": "if an advantage [for the better off] has the effect of raising the expectations of the lowest position, it raises the expectation of all positions in between" (TJ 80/69 rev.).

xxiv. Richard Musgrave, "Maximin, Uncertainty, and the Leisure Trade-Off", Quarterly Journal of Economics 88, 1974, 625-632, at p.632: "Implementation of maximin thus leads to a redistributive system that, among individuals with equal earnings ability, favors those with a high

preference for leisure. It is to the advantage of recluses, saints, and (nonconsulting) scholars who earn but little and hence will not have to contribute greatly to redistribution."

xxv. See e.g. Philippe Van Parijs, 1988. "Rawls face aux libertariens", in Individu et justice sociale. Autour de John Rawls (Paris: Le Seuil, 1988), 193-218; reprinted in P. Van Parijs, Qu'est-ce qu'une société juste? (Paris: Le Seuil, 1991) and Steven E. Byrne, 1993. A Rawlsian Argument for Basic Income (University College Dublin: Department of Politics), M.A. thesis, August 1993. The possibility of such a Rawlsian justification of an unconditional basic income is challenged by Eugene V. Torisky, "Van Parijs, Rawls, and Unconditional Basic Income", Analysis 53 (1993), 289-297; Magali Prats, "L'allocation universelle à l'épreuve de Théorie de la justice", in Documents pour l'enseignement économique et social 106, décembre 1996, 71-110; Georges Langis, "Allocation universelle et justice sociale", in Cahiers de Droit 37 (1996), 1037-1051; Colin Farrelly, "Justice and a Citizen's Basic Income", in Journal of Applied Philosophy 16 (1999), 283-96; François Blais, "Loisir, travail et réciprocité. Une justification rawlsienne de l'allocation universelle est-elle possible?", in Society and Leisure 22 (1999), 337-53.

xxvi. "While the notion of leisure seems to me to call for clarification, there may be good reasons for including it among the primary goods and therefore in the index as Musgrave proposes." A Reply to Alexander and Musgrave, (1974) in CP, 232-53 at p.253; and

I shall only comment here that twenty-four hours less a standard working day might be included in the index as leisure. Those who were unwilling to work under conditions where there is much work that needs to be done (I assume that positions and jobs are not scarce or rationed) would have extra leisure stipulated as equal to the index of the least advantaged. So those who surf all day off

Malibu must find a way to support themselves and would not be entitled to public funds. (See Rawls, PL 181-2 n9, modified version of "The Priority of Right and Ideas of the Good", (1988), in CP, 449-72, at p. 455 n7.)

xxvii. This conclusion is in broad agreement with Schaller (ARawls, the Difference Principle, and Economic Inequality,≡ op.cit.), who argues that the difference principle provides a strong case for a package of policies X possibly including both general employment subsidies for low-paid work and a low universal demogrant (or basic income) for all adults X that would "ensure that all working-age persons have an opportunity for at least an above-poverty income without thereby creating significant work disincentives" (p. 379). "If it turned out that the least advantaged representative persons were made worse off because too many people chose to live off their demogrant, then the demogrant would need to be reduced (or even eliminated). But [...] nothing in Rawls=s theory appears to prohibit conferring unearned or >undeserved= benefits on individuals like Smith [who opt for the life of a Bohemian artist]; the only question is whether doing so is in the best interest of the least advantaged representative person." (pp. 380-1) The fact that the best interpretation of the difference principle does not yield a firm justification for an unconditional basic income does not imply that no such justification is possible in a liberal-egalitarian framework. In Real Freedom for All (Oxford: Oxford University Press, 1995), I claim to supply one on the basis of a different understanding of maximin opportunities.

xxviii. Rawls, Justice as Fairness. A Restatement, (Cambridge, Mass.: Harvard University

Press, 2000) sect. 21.3

xxix. Here is a much earlier statement of a closely analogous idea: "Every man who possesses natural talents should know that if he shares them with his brothers in need, he possesses them rightly and in accordance with the will, intention and disposition of nature itself. Otherwise he is only a thief and a hoarder indicted by natural law because he holds onto what nature did not create exclusively for him." (Juan Luis Vives, De Subventione Pauperum [1526]. French translation: De l'Assistance aux pauvres (Bruxelles: Valero & fils, 1943), p. 141).

xxx. In this paragraph, I am assuming that making liberties effective (e.g. guaranteeing people's physical security in the streets through adequate policing) is distinct from insuring their fair value (e.g. by making sure people are sufficiently fed to be able to walk in the streets). Whereas the latter is governed by the difference principle (except for the Aworth of political liberties \cong), the former is covered by the first one. The high cost of making liberties effective for all (irrespective of securing their fair worth to all) is persuasively illustrated in Douglas Rae, "Democratic Liberty and the Tyrannies of Place," in Democracy's Edges, edited by Ian Shapiro, and Casiano Hacker-Cordón (Cambridge: Cambridge University Press, 1999), 165-192.

xxxii. An analogous mild notion of priority could be used to re-define the priority given to the least advantaged by the difference principle itself. This may provide a way of bringing the heavily handicapped under the difference principle without being driven into an "unreasonable" conception of justice, thereby making it possible to get rid of the rather ad hoc restriction of the difference principle to people "within the normal range."

xxxiii "Reply to Alexander and Musgrave" in Collected Papers, 232-53, at p.250.

xxxiii. These principles for individual conduct can be either "natural duties," which apply to everyone, such as the duty "to support and to comply with just institutions that exist and apply to us" (TJ 115/99 rev.), or obligations specific to particular positions, such as an obligation of noblesse oblige applying to "those who, being better situated, have advanced their aim within the system" (TJ 116/100 rev.).

xxxiv See G.A. Cohen, "Incentives, Inequality and Community"; AWhere the Action Is≡; and most recently If You're an Egalitarian, How Come You Are So Rich? (Cambridge Mass.: Harvard University Press, 2000); partly anticipated by Thomas C. Grey, "The First Virtue," Stanford Law Review 25 (1972): 286-327; Jan Narveson, "A Puzzle about Economic Justice in Rawls's Theory≡; Joseph H. Carens, "Rights and Duties in an Egalitarian Society," Political Theory 14 (1986): 31-49; and Barry, Theories of Justice, op.cit., Appendix C; and critically discussed by, among others, Philippe Van Parijs, "Rawlsians, Christians and Patriots. Maximin Justice and Individual Ethics", European Journal of Philosophy 1 (1993): 309-42; David Estlund, "Liberalism, Equality and Fraternity in Cohen's Critique of Rawls," Journal of Political Philosophy 6 (1998); Liam Murphy, "Institutions and the Demands of Justice", Philosophy and Public Affairs 27 (1998), 251-291; Paul Smith, "Incentives and Justice: G.A. Cohen's Egalitarian Critique of Rawls," Social Theory and Practice 24 (1998), 205-235; Andrew Williams, "Incentives, Inequality and Publicity", in Philosophy and Public Affairs 27.3 (1998), 226-248.; and "In Tax We Trust. Cohen on the Site of Distributive Justice". University of Warwick : Department of Politics; Frank Vandebroucke, Social Justice and Individual Ethics in an Open Society (Berlin & New York: Springer, 2000), etc.

xxxv Justice as Fairness. A Restatement, ə55

xxxvi Justice as Fairness. A Restatement, ə17.1

xxxvii Theories of Justice (Hemel-Hempstead: Harvester-Wheatsheaf, 1989), 399-400.

xxxviii Justice as Fairness. A Restatement, ə49

xxxix Ibid., ə 16.3, (Rawls's emphasis).

xl Ibid., see also TJ 87/75-6 rev.

xli Ibid., ə5.