How Self-less is the Liberal Subject?

Michael Sandel's Critique of John Rawls.

Michael Rosen,
Merton College,
Oxford.

In the vast literature which has developed around John Rawls's A Theory of Justice[1] there have been too few attempts to deal with the work as a sustained philosophical whole. For this reason alone - but also for the lucidity and seriousness with which it approaches its task - Michael Sandel's Liberalism and the Limits of Justice[2] is an important new contribution to the debate. Undoubtedly, one cause of this relative neglect is A Theory of Justice's sheer scale. But a more significant reason, surely, is that, like other revolutionary works, A Theory of Justice has so far altered the contours of the landscape it entered that the novelty of its project is no longer evident. In putting forward a fully worked-out liberal theory of political obligation at the the time he did, Rawls presented a challenge to the widely prevailing scepticism about the discipline of political theory.

One may identify the discipline of political theory by the central question which it asks. Just as epistemology asks: What can I know?, or moral philosophy: What ought I to do?, so political theory - to take Isaiah Berlin's formulation - gives an answer to the question: Why should anyone obey anyone else?[3] Now it follows, if we identify political theory with a question like Berlin's, that there are two ways in which the existence of a distinct discipline of political theory might be disputed. One might, first of all, for reasons sceptical or relativist, deny that the question can be given any useful answer. Alternatively, one might deny that the question of political obligation is a question which can be treated separately at all; that, although significant, it can only be answered in the

context of a comprehensive, shared moral framework. On this second view, it does not make sense to ask the question of political obligation except as a subordinate part of an objective theory of the good. Both of these denials are, of course, themselves significant parts of the tradition of Western political theory; what they deny, however, is political theory's independent validity.

One way of characterizing the liberal tradition in political theory broadly understood - the tradition whose founding fathers are Hobbes and Locke - is precisely as the attempt (or series of attempts) to maintain the idea of the autonomy of political theory in spite of these objections: to give a substantive answer to the question of political obligation without, at the same time, presupposing a shared, objective world of religion and morals. They were engaged in what Sandel calls (in a phrase echoing Rawls) the 'deontological project': the effort to develop political theory so that the right will take priority over the good.

And it is in this context that the force of Rawls's undertaking becomes apparent. It is nothing less than the attempt to revive the liberal project of political theory while, at the same time, doing justice to those critiques of moral realism which Anglo-Saxon philosophers found so persuasive in the post-war years. In this respect - as in very many others - Rawls's intentions are Kantian. He would like to detach the question of objectivity from that of realism; to show that one does not have to suppose a realm of independent moral facts to settle matters, in order for political theory to be a rational

and objective activity.

In reworking the idea of objectivity Rawls is led to propose a new method for political theory, the method of 'reflective equilibrium'. Reflective equilibrium is the attempt to bring into balance 'first-order' moral judgements (that such-and-such an action is wicked, or another laudable) and general principles. Its most revolutionary feature is this: political theory, as Rawls presents it, is not a construction upwards from a firm base so much as a movement outwards from an indistinct centre. Its technique is one of division and - through the confrontation of opposed elements - clarification. It is not, one might say, a question of finding the first turtle, but rather of determining where the elephants end and the turtles begin.

This conception implies a more limited ambition for political theory than as providing once-and-for-all answers to eternal questions. Rawls's concept of objectivity, like Kant's, has made considerable concessions to subjectivism. In particular, the objectivity of Rawls's conclusions is dependent, in the last analysis, on a coincidence between author and reader. He must suppose that the reader's deepest convictions will match - or, at least, prove sufficiently amenable to - his own. If they do not, then, by the assumptions of the method, the limit has been reached; reflective equilibrium cannot step outside the system of convictions to prove them wrong as a whole.[4]

The nature of Rawls's method suggests that, in principle, he will meet two kinds of criticism. There will be challenges to individual arguments - objections to sweeps of the shuttle

between particular judgement and general principle - and there will be more general refusals of assent when appeal is made to supposedly shared underlying assumptions. Yet, on their own, neither of the two kinds of criticism promises to be of much independent philosophical interest; the result, one suspects, will be either a nit-picking search for contradictions or else a simple registration of divergent intuitions.

According to Sandel, however, it is possible to combine the two aspects with conclusions which are much more far-reaching; into a critique revealing the foundations, superficially plausible but profoundly inconsistent, not just of the Rawlsian system, but of the entire deontological project. Thus, although Sandel (like many other commentators, conservative and radical, before him) dissents from the individualism which permeates Rawls's enterprise, this does not lead him to part company from it at the outset. Sandel regards Rawls's individualism as something more than an initial premise, simply to be accepted or rejected; he believes that it introduces tensions whose presence can be traced throughout the work. It is this perspective which gives Sandel's critique its focus - a focus which is missing from commentators who are either too sympathetic to stand back from individual arguments, or too critical to engage the arguments in detail. Nevertheless, his critique can be shown to fail.

Simply stated, Sandel sees Rawls as even more of a Kantian than he admits. Although Rawls claims to reject Kant's transcendental notion of the self, A Theory of Justice, Sandel argues, rests on just such a concept, for there is a necessary

connection between the deontological project and the abstract concept of the self. Sandel's claim is both substantive (that Rawls requires the abstract conception of the self in order to be able to reach his conclusions) and textual (that Rawls's text allows his commitment to the abstract self to show through).

I shall argue, however, that Sandel is wrong on both counts. Although I agree that the abstract conception of the self, as Sandel describes it, is unacceptable, I believe that Rawls neither needs it for his argument, nor commits himself to it in his text.

Initially, Sandel makes a division of conceptions of the self between two extreme poles, which he attributes (not, perhaps, without a degree of simplification) to Hume and Kant respectively. At the one end we have the 'radically situated' subject: a self without a focus which, having no distance between itself and the experiences which it undergoes, must acknowledge everything - all of its desires and emotions - as equally a part of itself. At the other, there is the abstract self, the 'subject of possession'. Here the self is detached from its attributes. They are things it <u>has</u>, not what it <u>is</u>. This self is defined, not by the possession of any attribute or group of attributes, but in its very power of possession: the capacity to have and, most importantly, to distance itself from and choose between desires and values.

Now, as Sandel himself points out, for Rawls 'a viable Kantian conception of justice' implies that 'the force and content must be detached from its background in transcendental

idealism'[5]. Nevertheless, in Sandel's view, this detachment is incomplete; Rawls retains a commitment to just the kind of conception of the self as a mere subject of possession which is at the heart of transcendental idealism.

What would bring Rawls to such an implausible doctrine? Sandel's answer focuses on Rawls's central argument, the 'original position'. The original position is put forward as an exercise in reflective equilibrium. Rawls asks us to compare our intuitions about justice with the principles which would be adopted by a group of individuals lacking information about their position in society and, crucially, about their conception of the good. Sandel gives two kinds of reason why a commitment to the abstract, Kantian self is required: reasons which stem from the need to make the adoption of the original position plausible, and reasons which stem from the need to generate the desired distributive conclusions.

Sandel's first arguments relate to Rawls's idea of the 'circumstances of justice'. To say that a society requires justice - some basis by which it is decided what 'giving each his due' amounts to - is to presuppose that certain general circumstances obtain. In the first place, there must be goods (broadly construed) to distribute and protect. Nor can there be such an abundance of them that no-one cares about their distribution and protection. There must, too, be some reason to be concerned to have explicit principles; distribution and protection do not just happen automatically and spontaneously. Now the fact that human societies have these characteristics appears to be an empirical generalization; and Rawls takes

himself to be free to make use of just such generalizations in the course of moral theory. Nevertheless, Sandel objects:

> But an empirical understanding of the original position seems deeply at odds with deontological claims. For if justice depends for its virtue on certain empirical preconditions, it is unclear how its priority could unconditionally be affirmed. Rawls says that he borrows his account of the circumstances of justice from Hume. But Hume's circumstances cannot support the priority of right in the deontological sense. They are after all empirical conditions. To establish the primacy of justice in the categorical sense Rawls' claim requires he would have to show not only that the circumstances of justice prevail in all societies, but that they prevail to such an extent that the virtue of justice is always more fully or extensively engaged than any other virtue. Otherwise, he would be entitled to conclude only that justice is the first virtue of certain kinds of societies, namely those where conditions are such that the resolution of conflicting claims among mutually disinterested parties is the most pressing social priority.
>
> (Sandel, p.30)

Sandel's objection fails, however, for what he offers against Rawls is a mixture of two arguments - neither of them persuasive.

Sandel's first move is to object in general to the existence

of empirical conditions ('if justice depends ... on ... empirical preconditions, it is unclear how its priority could unconditionally be affirmed') for something which is supposed to be unconditional. If dependent on conditions, how, then, unconditional? But this - as any student of Kant will realize - is a superficial objection. When the Kantian affirms that some action is an unconditional duty (say, that Jack returns the diamonds to Jill) he does not mean to say that, in any conceivable circumstances, Jack should give Jill the diamonds. What if Jill had stolen them? When the Kantian prescribes an action unconditionally he means that it is obligatory in these circumstances. What is meant in calling it unconditional is that the action ought to be performed whatever the state of motivation - whatever the desires, emotions and ends - of the agent called upon to do it. Returning to Rawls, then, we can see that it is not of itself an objection to the priority of justice that its existence requires empirical conditions; it is just in those circumstances that justice has its priority.

At this point, though, Sandel's objection changes. No longer is it objectionable that the circumstances of justice should be an empirical matter at all, but that, as a matter of fact, justice cannot universally be taken to be prior. Even where the circumstances of justice obtain, it is not necessarily the case, Sandel says, that 'justice is always more fully or more extensively engaged than any other virtue' or that 'the resolution of conflicting claims among mutually disinterested parties is the most pressing social priority'.

But, clearly enough, to take Rawls's claim for the priority

of justice in this way is to change it drastically. To say that justice is the first virtue of social institutions is not to say that it is 'more extensively engaged than any other virtue' or that it is 'the most pressing social priority'. Virtues are not like that; priority is not equivalent to strength or urgency. Rawls's own analogy explains this very well. He writes: 'Justice is the first virtue of social institutions as truth is of systems of thought'. (Rawls, p.3) Yet truth is not something we weigh up against other intellectual qualities. Truth is not even, always, the virtue of systems of thought which it is reasonable to make our most pressing concern (should we prefer a thought-system which tells us only true things we already know to one which is informative but sometimes makes mistakes?) In no way does this undermine the priority of truth, for its precedence is a logical one. And so it is, Rawls is able to claim, with justice. Thus Sandel's objection fails.

It is worth noting the structure of Sandel's case. The claim is that Rawls's argument, as presented, is inadequate. Thus we must look for a less obvious, latent argument, and this will, the implication is, lead us to the Kantian view of the self. But if, as I believe, Sandel's initial objections can be shown to be unconvincing, the need to attribute an implicit, transcendental Geheimtext evaporates.

However, Sandel's argument is not only substantive. There is also, he maintains, an explicit commitment on Rawls's part to the abstract view of the self. Sandel's crucial piece of

evidence - the phrase is one he repeats throughout his book - is Rawls's statement that 'the self is prior to the ends which are affirmed by it'. Since the sense of this priority is vital to the argument, it is worth quoting in full the passage in which appears:

> ...the structure of teleological doctrines is radically misconceived: from the start they relate the right and the good in the wrong way. We should not attempt to give form to our life by first looking to the good independently defined. It is not our aims that primarily reveal our nature but rather the principles that we would acknowledge to govern the background conditions under which these aims are to be formed and the manner in which they are to be pursued. For the self is prior to the ends pursued by it; even a dominant end must be chosen from among numerous possibilities. (There is no way to get beyond deliberative rationality.) We should therefore reverse the relation between the right and the good proposed by teleological doctrines and view the right as prior. The moral theory is then developed by working in the opposite direction.
>
> (Rawls, p.560. Quoted, Sandel, p.19. Sandel's emphasis; sentence in brackets omitted by Sandel.)

Sandel, it will be recalled, locates conceptions of the self between two poles: the 'radically situated' Humean self, for which all of its attributes are equally essential, and the

radically distanced, Kantian self, for which what is solely essential is the capacity to choose. Sandel believes that the passage just quoted carries a commitment to the Kantian extreme. But this is not so. The thrust of this passage is, indeed, directed against the kind of radically situated subject which Sandel finds objectionable, but Rawls's polemic does not carry him to the other, pseudo-Kantian extreme. To the contrary, there are features of Rawls's conception of the self - most especially in his understanding of the way in which individuals choose coherent plans of life - which are quite at odds with the picture of the abstract choosing self ascribed to him.

The quoted passage is offered by Rawls as a concluding summary of his argument against hedonism. Its content is easily recapitulated. Hedonism, Rawls takes to be the belief in pleasure as the dominant end. It is, he says, either trivial or false. If we take the pleasurable to be identified by reference to what we in fact choose (the saint chooses martyrdom because it gives him greater pleasure, and so on) then the theory is circular. But, if pleasure is independently defined as a psychological state, then it is simply not a reasonable dominant end. As Rawls says:

> We need only note that once pleasure is conceived, as it must be, in a sufficiently definite way so that its intensity and duration can enter into an agent's calculations, then it is no longer plausible that it should be taken as the sole rational aim.
>
> (Rawls, pp. 556-7.)

But by no means does this commit Rawls to an abstract view

of the self. Rawls rejects a conception of the good as following the dominant end of the desire for pleasure because he rejects the identification of the self with the particular set of desires it may have at any one time. But it does not follow that the self is radically distanced from or entirely independent of its desires. Simply, Rawls takes it to be an evident fact that we often have to adjudicate between different sets of desires - and that the way that we do this will not be a matter of moral indifference to us: 'we can choose now which desires we shall have at a future time'. (Rawls, p.415) This is not to say, however, that our choice takes place in an existentialist vacuum. As Rawls puts it:

> At any given time rational persons decide between plans of action in view of their situation and beliefs, all in conjunction with their present major desires and the principles of rational choice. Thus we choose between future desires in the light of our existing desires, including among these the desire to act on rational principles
>
> (Rawls, p.415, my emphasis)

Rawls is just locating his conception of the self somewhere in the spectrum between Kant and Hume. Our present aims, beliefs and desires are not inessential in the sense of being dispensable to us; it is only in terms of them that we can decide between competing rational plans of action. But they are not essential in the sense that the self may be fixed by reference to any single group of them, once and for all. It is in this way - that it is not to be reduced to its ends, not that

it is metaphysically abstracted from them - that Rawls's self is 'prior to the ends affirmed by it'.

To accept this dispels a great deal of hot air. Rawls's self, says Sandel, is 'shorn of all contingently-given attributes, [it] assumes a kind of supra-empirical status, essentially unencumbered, bounded in advance and given prior to its ends, a pure subject of agency and possession, ultimately thin'. (Sandel, p.94.) But it is plainly false to suppose that any epistemology of the self which refuses (rightly) to reduce the self to its immediate desires, is thereby committed to treating the self as the kind of metaphysical phantom that Sandel's rhetoric projects.

So far, then, Sandel's arguments fail to convince. But it is fair to say, the argument for a transcendental derivation of the circumstances of justice, and the textual claim of the self's priority over its ends, are only intended to reinforce the central part of his case: that Rawls requires the abstract conception of the self in order to render plausible the derivation of his distributive principles in the original position. As Sandel puts it, the abstract self makes possible Rawls's passage from libertarianism to egalitarianism.[6]

Sandel's interpretation draws on an argument of Robert Nozick's. What, ask Nozick and Sandel, allows Rawls to treat assets in the original position as if they were assets of the community - a common 'cake' to be sliced up as the contracting parties think best? Why, in particular, do we not regard assets

as, from the very beginning, owned by individuals? If assets are owned by individuals, then, what the parties in the original position are doing is not dividing up a common resource but redistributing: taking from one group of individuals to give to another. Rawls's justification for treating assets as common, says Sandel, lies in the abstract notion of the self; because whatever is empirical is inessential to the self, one need not suppose that any individual is especially entitled to possess or benefit from empirical assets. On this interpretation, we have no justification in claiming as our own the return which a free market would allow from our assets; for they are not, in that strong sense, ours.

Nonetheless, for Sandel, Rawls's argument is a failure. What the abstract conception of the self establishes, at best, is a 'no-ownership' account of natural assets - not that these can be regarded as the property of the community as a whole. Here Sandel and Nozick begin to move in opposite directions. For Nozick, since the individualist premises do not lead to egalitarian conclusions, the conclusions are to be rejected; for Sandel, it is the individualist premises which must be substituted:

> Rawls might deny that the difference principle uses me as a means to others' ends, not by claiming that my assets rather than person are being used, but instead by questioning the sense in which those who share in 'my' assets are properly described as 'others'. Where the first defense presses the distinction between the self and its attributes, the second qualifies the

> distinction between the self and the other by allowing that, in certain moral circumstances, the relevant description of the self may embrace more than a single empirically-individuated human being. The second defense ties the notion of common assets to the possibility of a common subject of possession.
>
> (Sandel, pp.79-80.)

Plainly, Sandel's argument is pulling in opposite directions at the same time. On the one hand he is arguing that Rawls needs the abstract conception of the self to underpin the move from libertarianism to egalitarianism. But, on the other hand, he claims that it fails to do so. I shall argue, however, that this conflict is illusory. Rawls's argument can be reconstructed so as to meet the Nozickian objection without presupposing either the abstract subject of possession which Sandel attributes to him or the collective one which Sandel proposes as a substitute. As Sandel presents it, the dispute about distributive justice refers us back to the metaphysics of the self. Each principle of distribution can be related to a different 'subject of possession'. Hence, to 'rescue' Rawls it is necessary to find the subject of possession appropriate to egalitarianism. It is from this assumption that I dissent.

Nozick objects to egalitarianism that its redistribution involves disposing over the legitimately held property of individuals. It is an objection which can be met in two stages. To start with, the egalitarian may question how the libertarian thinks that this property comes to be 'legitimately held' in the first place. The egalitarian is accused of putting individual

property into collective hands for the purposes of the original position. But, really, a very similar accusation could be framed in the other direction. Failing some further argument, what is to say that the property which is held individually is held justly? In his own book, Anarchy, State and Utopia,[7] Nozick does, it is true, answer this with his own account of the original appropriation of goods. He offers a Gedankenexperiment - corresponding to Rawls's original position - to show how property might originally have come to be legitimately held. It is a model based on Locke's Second Treatise; but, without Locke's theological and teleological premises, the Nozickian account is highly unconvincing.[8] Yet, it - or some equivalent account - is a crucial premise for the argument against the egalitarian. Without a convincing account of original ownership there is no more reason to take the question of distributive justice, with Nozick, to be: When, if ever, can we deprive one individual of property in favour of another?, than, with Rawls, (and, incidentally, Locke himself) as: How can individuals have claims on goods held by the community in common?

Thus, the first stage of the reply to Nozick is that, failing a justification of private possession, there is no reason to adopt the libertarian's perspective of a world of owned things rather than the egalitarian's picture of 'common assets'. But, surely, the libertarian will say, even if the Nozickian picture of how goods come to be acquired by individuals is rejected, the argument against Rawls as regards his treatment of natural assets remains. According to Rawls, even our natural assets - our bodies and our minds - can be treated by the community as common

resources (not, however, that Rawls thinks we can legitimately be deprived of them; but we have no rights to benefit from them). One argument for this - the argument that Sandel attributes to Rawls - is that even these natural characteristics are metaphysically, and, hence, morally, inessential to us. I think, however, that a more interesting and plausible argument can be developed.

Taken alone, of course, our natural assets are of limited value to us. True, our intellect and our bodies may give us some narcissistic pleasure, but we are social beings, and it is in depriving us of opportunities to make use of our assets in conjunction with others that Rawls, it is held, is disposing over them unjustly. Nozick's attitude is clear: it is as much a violation of my property to prevent me from using it to make free exchanges as it would be to deprive me of it directly. But must the egalitarian share this attitude? Freedom, as Sandel himself points out, is not the only moral principle involved in exchange; there is also fairness.

Society is, the egalitarian can argue, a system of <u>productive cooperation</u>. Individuals bring together a diversity of natural assets which, together, are transformed into a social product, more valuable than the original assets. The question, of course, is how this social product should then be distributed. To make distribution the outcome of an unrestricted market system would be one way. Assets would only be brought into productive combination by being acquired (for a price) by some individual, to whom, then, the social product belongs <u>in toto</u>. The intuition which makes this system appear equitable is that the

sale of the original assets will only take place if their owners receive a price which they consider fair. But this is a limited intuition, for there are a whole series of circumstances in which inequalities between negotiating individuals - what one may call in general inequalities in market power - will lead to outcomes which are not just. Relative scarcity and market power do not, in the egalitarian's view, provide a satisfactory basis for establishing equity in exchange.

The obvious alternative - that the product should be divided in proportion to the value of the original assets contributed - can be shown to be equally unsatisfactory, however. For by value is meant, presumably, magnitude of contribution, and there is, in principle, no way that this can be identified. The concept of marginal productivity is often suggested for the purpose. Marginal productivity values factors of production by establishing the difference between the quantity produced with, and the quantity produced without the last unit of the factor. But this is no way to measure contribution. The analogy with a chain explains succinctly why: take away any one link from the chain and it holds nothing. But from that it follows that each link is 'contributing' (marginally) the whole strength of the chain - which is obviously absurd. Contribution simply cannot be separated out in this marginalist way.

The egalitarian takes society to be, in an important sense, analogous to such a chain: each individual is unique and indispensable to the good of the whole[9]. Thus, each must be seen as having equal claims. And it is this which governs the legitimate interest society takes in the content of exchanges

between individuals.

Nozick's objections are met by the egalitarian standpoint outlined here: it justifies treating the community's goods as shared resources, and it refuses the libertarian intuition that there is a particular sanctity to the system of free exchange. It is clear, however, that Nozick's objections, strictly speaking, are deflected rather than refuted. It is shown that Rawls's egalitarian vision does not have to depend (as Nozick and Sandel allege) on a metaphysical view of the self as an abstract subject of possession - or, indeed, on any other metaphysical alternative. But, undeniably, it draws its force from a particular moral perspective - a vision of society as cooperative enterprise between unique and mutually indispensable individuals. In this sense, Nozick is right: the egalitarian conclusions of Rawls's difference principle depend upon the intuition which supports setting up the original position in the particular way Rawls does (all have a claim, all have a veto). But this should be no surprise. It is just what the method of reflective equilibrium would lead one to expect: the whole point is to make clear the consequences of certain general intuitions about justice. These intuitions may not be shared; the Judaeo-Christian picture of human beings as 'members one of another' may no longer have much force in Western society. But in that case the spade turns. It is not, as Sandel would have it, a matter of constructing a philosophical argument in favour of a collective subject; Rawls's theory of justice rests on a shared moral understanding, not idealist metaphysics.

Borrowing Nozick's argument does not, then, give Sandel the substantive conclusion he needs: that Rawls is forced into adopting an abstract notion of the self with no intrinsic connection to its attributes. Nevertheless, his critique is not completely disposed of without discussion of the striking interpretation he puts on Rawls's rejection of the idea of distribution according to desert. Rawls rejects the idea of distribution according to desert, Sandel says, because, for him individuals have no intrinsic worth:

> The principles of justice do not mention moral desert because, strictly speaking, no one can be said to deserve anything. Similarly, the reason people's entitlements are not proportional to nor dependent upon their intrinsic worth is that, on Rawls's view, people have no intrinsic worth, no worth that is intrinsic in the sense that it is theirs prior to or independent of or apart from what just institutions attribute to them.
>
> (Sandel, p.88, Sandel's emphasis)

This is a quite remarkable doctrine to attribute to a Kantian such as Rawls, and a Kantian analogy explains why it is misguided. When Rawls denies that the concept of moral worth can provide a first principle of distributive justice, his argument appears to run parallel to Kant. Kant's categorical imperative enjoins us to treat people as ends, not means. But, in that form, it fails to identify specific courses of action for us to take. What is it, in any circumstances, to treat someone as 'end' rather than 'means'? This is not something to be settled at the level of a priori principle. Yet the fact that

the idea of treating someone as an end is not sufficient to give a determinate content to ethics certainly does not mean that the categorical imperative fails to acknowledge the intrinsic worth of individuals; no more should Rawls's rejection of desert be taken as a sign that 'people have no intrinsic worth' for him.

It must be admitted, however, that Rawls's argument is a thin one. It does not seem conclusive against someone who argues (in the original position) for the principle of rewarding moral worth and of settling later what that worth amounts to - what people in fact deserve. Rawls does, though, offer other arguments. One is that our productive capacities are not, of themselves, virtues - our strength or intelligence may make us useful, but they are not signs of goodness. This is, of course, true, but, again, it is of limited force. Should not those qualities which genuinely do represent moral worth - for example, courage in pursuit of the good of the community - be rewarded, nonetheless? Rawls's counter-argument - his best and simplest, in my view - is to claim that this is to mistake the purpose of distributive justice. Distributive justice should just not be aimed at rewarding virtue. On religious grounds, one might say, the reward of virtue is a matter for God alone. Less religiously, one might add that the proper reward for virtue is private praise, not public profit. Rawls suggests that the belief in distribution according to virtue may be a product of a superficial symmetry between retributive and distributive justice: just because the former exists to punish the wicked, one should not conclude that it is the task of the latter to reward the virtuous. But, even if these considerations fail to

convince protagonists of distribution by desert,[10] the important point, for present purposes, is that they in no way depend on any kind of scepticism about the intrinsic worth of individuals.

This is reinforced when one compares Rawls's treatment of the idea of natural duty with Sandel's account of it. Justice, says Rawls, is a 'fundamental natural duty'; that is, it is one of the duties which 'hold between persons irrespective of their institutional relationships; they obtain between all as equal moral persons'. (Rawls, p.115) Sandel takes this to mean that natural duties are derived from a hypothetical contract:

> Notwithstanding their contractarian derivation, the natural duties apply without reference to our voluntary acts ...
>
> (Sandel, p.111)

Rawls is thus held to have both a contractarian political philosophy and a contractarian ethics. But this is to reverse Rawls's point. When he writes that '... even though the principles of natural duty are derived from a contractarian point of view, they do not presuppose an act of consent, express or tacit in order to apply' (Rawls, p.115) Sandel takes him to mean that all natural duties are derived from the contractarian point of view. To the contrary, the context in the discussion of justice makes it clear that what Rawls is claiming is that even where the principles of natural duty are derived from a contractarian basis (i.e. in the case of justice) their validity is not institutional. The morality of natural duty thus precedes the morality of contract.

But from where, then, does Rawls derive the principles of natural duty - our duty to help another, or not to be cruel? The answer is that they are not derived: their existence forms part of that shared framework to which the reflective equilibrium theorist must repeatedly appeal. This is unacceptable, according to Sandel, for it 'runs counter to the core of the deontological project'. (Sandel, p.116) What it runs counter to, in fact, is the misinterpretation of the deontological project's foundations.

We can see, then, that there is no force to Sandel's indictment. Its central contention, that Rawls is compelled to adopt an abstract conception of the self, fails, both substantively (it is not necessary for the claim of the primacy of justice, nor for the egalitarian derivation of the difference principle in the original position) and textually (it depends on a distortion of the sense in which, for Rawls, the self is prior to its ends; of his rejection of desert as a basis of distribution; and of his claim that justice is a natural duty).

I have argued that, for his theory, Rawls needs only a conception of the self placed upon the reasonable spectrum between the 'radically situated' and the 'abstract' self. But, throughout his book - not least in his references to the 'common subject of possession' - Sandel hints that, for him, the question of the nature of the self is less of a spectrum than a triangle. At the apex stands a collective self whose ends are not chosen; not just in the sense that they are not set arbitrarily or indifferently (in that sense the Rawlsian self does not choose

its ends, either) but in the sense that its ends are discovered as an objective moral good. From that perspective, of course, less ambitious conceptions of the self - even ones which do not lie at either the 'Humean' or 'Kantian' pole - will be unsatisfactory. But Sandel never goes beyond allusion; he never makes clear how such a conception of the good can be made determinate, or how it is to be discovered. Should he succeed in doing so, he will, of course, have made a major, positive contribution to moral theory. But, as things stand, the collective subject does not contribute to his critique. For it hardly seems a severe criticism of Rawls if the worst that can be said against him is that he is not Hegel!

NOTES

1. J. Rawls, A Theory of Justice (Oxford: Oxford University Press, 1972). Hereafter: Rawls.

2. M. Sandel, Liberalism and the Limits of Justice (Cambridge: Cambridge University Press, 1982). Hereafter: Sandel.

3. In 'Does Political Theory still Exist?', in P. Laslett and W. Runciman (ed.), Philosophy, Politics and Society, Second Series (Oxford: Basil Blackwell, 1962), pp. 1-33.

4. Although Rawls is not a moral realist, it does not seem to me impossible that a moral realist should adopt Rawls's procedure. In that case, reflective equilibrium - like the Platonic dialogue - would be a method for the recovery of an objective moral truth, latent in us all.

5. J. Rawls, 'The Basic Structure as Subject', American Philosophical Quarterly, Vol.14, No. 2, 1977, pp. 159-65, p.165. Quoted, Sandel, p.13.

6. Like Sandel, I take Rawls's difference principle to be highly egalitarian. In the limiting case - where the magnitude of social welfare is independent of the mode of distribution adopted - the difference principle prescribes equality.

7. R. Nozick, Anarchy, State and Utopia (New York, N.Y.: Basic, 1974).

8. Obvious problems are: how to identify the scope of a single act of appropriation; how to proceed when there is no longer 'as much and as good' left for others. At this point, I take Locke to make use of the 'fundamental law of nature', that everything is to be preserved, as much as may be. This course is not open to Nozick.

9. This is, of course, a reconstruction of Rawls; but one which, I maintain, is fully in the spirit of A Theory of Justice. Rawls emphasizes the cooperative nature of society and says that the difference principle expresses the ideal of fraternity. His reflections on the idea of community (in contrast to the Marxist tradition) ascribe a positive role to the division of labour: each individual comes to be seen as having unique and indispensable qualities. There is a striking resemblance on this question between Rawls's attitude and Durkheim's (see, The Division of Labour in Society (New York, N.Y.: The Free Press, 1964))

10. For a forceful advocacy see J.R. Lucas, On Justice (Oxford: Oxford University Press, 1980).