Rawlsian Justice and Preferences Over Work

by

Wylie Bradford

wbradfor@efs.mq.edu.au

Abstract

Rawls’s assertion that the specification of property rights over the means of production is not determined by the theory of justice is shown to rest on shaky foundations. His account of the subjective circumstances of justice, which serve as an empirical constraint on the deliberations of those in the original position, is shown to reflect neoclassical economic theory in which preferences over work are excluded from welfare analysis. As a result, those in the original position are concerned only with the distributive outcomes of the basic structure, and not at all with its form in terms of the modalities of ownership of productive assets. This in turn is shown to conflict with Rawls’s professed methodology, in which constraints are not to be placed upon the conceptions of the good held by those in the original position. There is, however, very little scope for Rawls to amend his argument on this point without compromising the integrity of the whole. Hence, there is no good reason to exclude property rights over the means of production from consideration in the liberal theory of justice.

JEL Classification: A13, D63, P16
Introduction*

An important aspect of any discussion of economic justice is the question of property rights over the means of production. The significance of the issue follows from the obvious importance of such relations for the material productivity of any given society, and for the social relations that will emerge from the economic interactions of its citizens. Given the undeniable complexity of the issues involved, it is surprising that Rawls provides a disarmingly simple answer, one that survives through the considerable transformation that takes place in his overall theory of justice.

For Rawls, the theory of justice provides no guidance on the question of whether private ownership of the means of production should be one of the institutions included in the basic structure of the just society (1971, henceforth TJ, p280). Instead, the principles of justice are to be derived separately, treating only as basic the right to personal property, leaving the specification of property rights over the means of production to be determined by the history and traditions of the society in question. This feature of Rawls's argument, whereby history is summarily accorded priority in terms of shaping the form of the basic structure, and is apparently immune to analysis in terms of the principles of justice, is referred to as historical dominance by Bradford (2000a).

An obvious implication of historical dominance is that the parties in the original position are not concerned with the form of the basic structure (at least in terms of its ‘economic’ aspect), but rather only with its distributive outcomes. That is to say, they are not concerned whether a form of either capitalism or socialism obtains. This reflects the fact that, contra rights to personal property, property rights over the means of production are not considered basic. Hence, they need not receive the constitutional protection implied by the first principle of justice as fairness.

In what follows, it is shown that Rawls’s argument rests, in part, on his reliance on neoclassical economic theory, as represented by Koopmans (1957). In particular, it is the assumptions concerning preferences over work (i.e. relating to different production processes) underlying that approach that leads Rawls to affirm the conditional comparability of outcomes, the claim that any specific desired welfare outcome is achievable under any assignment of property rights over the means of production. As the parties in the original position concern themselves only with the outcomes of the basic structure, it follows that the conditional comparability of outcomes suffices to render them uninterested in the specification of property rights over productive assets.

However, it is demonstrated that the irrelevance of the form of the basic structure in Rawls’s argument is intimately related to his account of the circumstances of justice, particularly the subjective circumstances. The circumstances of justice are general facts, known to those behind the veil of ignorance, the function of which is to ensure that the conjectures of those in the original position remained tied to fundamental aspects of the reality to which they are to be applied. Rawls’s account of the subjective circumstances of justice is shown to be, in effect, a restatement of the neoclassical position regarding preferences over work; the circumstances of justice, in other words, are theoretical inferences rather than facts.

Thus, the exclusive focus on the outcomes of the basic structure stems from assumptions made concerning the values of the parties to the deliberations in the

---

* The author acknowledges the benefit of comments from Geoff Harcourt, Ken Tsutsumibayashi, Ugo Pagano and Rod O’Donnell. All are resolved of responsibility for any remaining errors.

1 This is despite the fact that he acknowledges that the arrangement of the background institutions of property is one of the “moral problems of political economy” (TJ p265)
original position. This is inconsistent with Rawls’s insistence that no strong assumptions are made concerning the conceptions of the good held by the agents in the original position. Hence, Rawls can avoid consideration of property rights over the means of production only by ignoring his own methodological precepts. Furthermore, it is shown that the difficulties caused by Rawls’s account of the subjective circumstances of justice cannot be avoided without significant modifications of his central argument.

The Circumstances of Justice

In Section 40 of TJ, Rawls suggests an interpretation of the original position that is intended to imbue the device with more philosophical weight than it would have as a simple thought experiment. To this end he invokes Kant on the appropriate conception and bases of moral philosophy, and relates that his own approach is an attempt to conform to, and apply, Kantian principles.

Kant, according to Rawls, holds that moral philosophy is the study of the “conception and outcome” of a “suitably defined rational decision” (TJ p251). Moral principles are “legislation for a kingdom of ends” and furthermore must be “agreed to under conditions that characterize men as free and equal rational beings” (TJ p252). When persons choose principles of action that express their nature as free and equal rational beings they are acting autonomously. This is to be contrasted with heteronomous action, which stems from principles chosen on the basis of social position, natural endowments, the type of society in which the person lives or the specific things which he/she desires.

Rawls contends that the deliberative stage of the original position is a mechanism for realising Kant's conception of moral philosophy; the veil of ignorance prevents the parties from choosing heteronomous principles. As a result, the principles agreed to reflect the nature of persons as free and equal rational beings, and thus permit autonomous action.

Furthermore, the principles of justice derived in the deliberative stage of the original position are categorical imperatives in Kant's sense. These are rules of conduct which apply to a person “in virtue of his nature as a free and equal rational being” (TJ p253). These are to be contrasted with hypothetical imperatives, which “[direct] us to take certain steps as effective means to achieve a specific end”, where the desire for the end involved is not a necessary condition for being a rational human being (ibid).

Rawls asserts that the structure of the original position ensures that the principles of justice are categorical imperatives, as the argument for the principles is not based on the assumption that the parties have any particular ends. Instead they desire only the primary goods. This desire, based on the ‘thin theory of the good’, is part of the definition of rationality; as primary goods are all-purpose means to any end whatever, it is rational to want them regardless of other desires. Thus,

2 Strictly speaking, Rawls’s focus in this case is on the first stage of the four-stage sequence that makes up the original position. This first stage is denoted the deliberative stage by Bradford (2000a). The purpose of the deliberative stage is the derivation of the principles of justice – it is here that the ‘veil of ignorance’ is at its thickest. Rawls labels the subsequent stages of the original position the constitutional, legislative and judicial stages. He confuses the reader somewhat by continuing to use the term ‘original position’ to denote what is here called the deliberative stage, even after he has introduced the four-stage sequence as an elaboration of the original position device.

3 The ‘thin theory of the good’ is ‘thin’ in the sense that no assumptions are made concerning what individuals actually value, or ought to value. That is, they know that they possess a conception of the good without knowing to what it refers. The thin theory of the good is examined in more detail.
to act from the principles of justice is to act from categorical imperatives in the sense that they apply to us whatever in particular our aims are. This simply reflects the fact that no such contingencies appear as premises in their derivation (TJ p253).

Finally, Rawls contends that the deliberative stage can be interpreted as the viewpoint from which the noumenal self sees the world. It is at this point that he introduces a significant modification of the Kantian project. An obvious problem with viewing the parties in the original position as the collection of noumenal selves is that the latter are transcendental, lacking in empirical foundation. Sandel (1982 p13) notes that it is unclear how...abstract, disembodied [subjects] could without arbitrariness produce determinate principles of justice, or how in any case the legislation of such...[subjects] would apply to actual human beings in the phenomenal world.

Rawls’s response is to 'constrain' the hypothetical choice situation, from which the principles of justice are derived, in such a way as to ensure a determinate outcome applicable to actual human societies. He achieves this by allowing a minimal amount of empirical information to be available in the deliberative stage, namely, that the parties are aware that they are subject to the circumstances of justice (TJ pp. 126-128). The circumstances of justice comprise both the subjective and the objective. The aspect of the former stressed by Rawls in TJ is that the members of society are mutually disinterested; each has his or her own conception of the good, which they consider worthy of recognition and deserving of satisfaction, but they take no interest in the interests of others. Thus there are competing claims made on the benefits of social subsequently.

4 The noumenal self in Kant is the self qua subject of experience, the transcendental self, given prior to its ends. As Sandel (1982 p8) has it

This something further [beyond the stream of sense data], which we cannot know empirically but must nonetheless presuppose as the condition of knowing anything at all, is the subject itself. The subject is the something 'back there', antecedent to any particular experience, that unifies our diverse perceptions and holds them together in a single consciousness. It provides the principle of unity without which our self-perceptions would be nothing more than a stream of disconnected and everchanging representations, the perceptions of no one. And while we cannot grasp this principle empirically, we must presume its validity if we are to make sense of self-knowledge at all.

The converse of the noumenal self is the phenomenal self: the self qua object of experience, which can be empirically known and is subject to all the laws of nature and regularities of cause and effect.

5 Rawls claims that

The theory of justice...tries to present a natural procedural rendering of Kant's conception of the kingdom of ends, and of the notions of autonomy and the categorical imperative. In this way the underlying structure of Kant's doctrine is detached from its metaphysical surroundings so that it can be seen more clearly and presented relatively free from objection (TJ p264).

6 As Rawls is committed to making no restrictive assumptions about the parties' conceptions of the good, he insists that the interests of individuals need not be egoistic, but could involve “ties of sentiment and affection” which would become known once the veil of ignorance is lifted (TJ p129). It is interesting to note Rawls's strange definition of egoism, and how it relates to the concept of primary
cooperation. Rawls argues that if this were not the case, if society was “an association of saints agreeing on a common ideal”, then justice would be nugatory (TJ p129). It is the fact that there is conflict over the distribution of primary goods which makes justice necessary, and establishes its primacy as a virtue.\textsuperscript{7}

In his later work, Rawls (1993, henceforth PL) moves away from Kantian constructivism in favour of what he calls political constructivism. The essence of the change in perspective is that the principles of justice are no longer taken to represent a comprehensive moral doctrine shared by all citizens of the well-ordered society. Rather, they are a set of principles that are modular in character, capable of being affirmed by citizens holding a wide range of comprehensive doctrines without being seen to derive their legitimacy uniquely from any particular member of that set. As such, the case for the principles cannot be based, as in TJ, on the fundamental nature of human beings, an approach Rawls considers appropriate for a moral theory. Instead, the foundations of justice as fairness must be political in nature, reflecting the reality that modern democratic states are characterised by a plurality of comprehensive doctrines.

Thus, in PL, Rawls defines the subjective circumstances of justice (SCJ) as the existence of pluralism as such, delimited to reasonable pluralism in the case of a well-ordered society of justice as fairness (p66). It is the awareness of the need to produce principles of justice that are compatible with this plurality of extant basic positions that represents the empirical constraint on the deliberations of the parties in the original position.

Although this account is superficially different from that provided in TJ, it is apparent that beneath the fact of pluralism lies the reality of competing claims on resources and output. For instance, those moral (or religious) doctrines that hold that other sentient creatures ought not be utilised as food for humans imply very different usage patterns for land and other natural resources to those in which meat-eating is not proscribed. Conflicting moral positions on abortion or euthanasia can likewise be represented in terms of their differing claims regarding the allocation of health and medical resources, and the form of the output resulting from such allocations. On a grander scale, religious doctrines that differ on their attitudes to the moral worth of material gain, and ethical systems that weigh competitive and cooperative behaviour differently, will involve starkly contrasting distributive claims. Indeed, a society in which pluralism did not involve competing claims over resources and output would be goods. The primary goods are all-purpose means to any ends that may figure in a conception of the good. As such, it is implied that they cannot be desired for their own sake. As wealth and income, and powers and opportunities, figure among the primary goods it follows that these will not be ends pursued by any person. Now even casual observation of most human societies seems to make nonsense of this claim. The reasoning underlying this aspect of Rawls's argument is his definition of egoism. A conception of the good is considered to be egoistic if it treats “wealth, position, and influence, and the accolades of social prestige” (ibid.) as ends. By assuming that parties in the original position are not egoistic, Rawls at one definitional stroke ensures both that no person in the just society will desire wealth, prestige and power for their own sake, and, more importantly, that the device of primary goods is meaningful. Despite Rawls's claim that he makes no restrictive assumptions about the parties' conceptions of the good, it appears that without the restrictions entailed in the definition of egoism the parties in the original position would have nothing to distribute. In other words, the argument would stall at the outset.

\textsuperscript{7} Sandel (1982) replies ingeniously to this that, given this reasoning, the removal of the circumstances which give rise to the need for justice must be a virtue of equivalent importance to justice itself (pp. 312). If the existence of danger makes physical courage a prized virtue, then the eradication of danger ought to be accorded the same importance as the inculcation of bravery. Rawls does not follow this route, and treats justice as strictly a remedial virtue. In this light, those who consider Rawls an apologist for mature American welfare capitalism are provided with telling ammunition.
perilously close to the “association of saints” in which discussions of justice have little import. Hence, other differences notwithstanding, the accounts of the SCJ in TJ and PL can be seen to share the feature that the presence of competing claims over the fruits of social cooperation represents a constraint to which the principles of justice formulated in the original position must conform.

However, the fact of competing claims is not sufficient to establish the primacy of justice as a virtue, and so Rawls marries it with the objective circumstances of justice (OCJ). The most important of these is the moderate scarcity condition. The parties in the original position know that resources are not limitless, and that “mutually advantageous arrangements”, while feasible, will not be so productive as to satisfy all demands (TJ p127).

The parties in the deliberative stage are aware that, beyond the veil of ignorance, the circumstances of justice hold. The only empirical information available to them, about the society in which they will live, is what they can infer from their knowledge that the circumstances of justice apply. Hence, although the parties in the deliberative stage have complete freedom to choose whatever principles of justice they desire, they also have a desire to express their nature as rational and equal members of the intelligible realm with precisely this liberty to choose, that is, as beings who can look at the world in this way and express this perspective in their life as members of society. They must decide, then, which principles when consciously followed and acted upon in everyday life will best manifest this freedom in their community, most fully reveal their independence from natural contingencies and social accident (TJ p255).

Rawls contends that justice as fairness, a “theory of human justice” which “counts among its premises…the elementary facts about persons and their place in nature”, is the conception of justice that would satisfy these conditions in the deliberative stage (TJ p257).

What is the connection between Rawls's use of the circumstances of justice as the empirical lifeline to the agents in the deliberative stage, and the central theme of this paper, the contention that in Rawls's system the form of the basic structure (particularly with respect to property rights in the means of production) is subordinate in importance to the distributive outcomes thereof? The answer lies in Rawls's understanding of the circumstances of justice. To him they are facts, the most elementary and general facts about persons and social reality.

However, it is clear upon reflection that the circumstances of justice are not facts at all, but are better thought of as assumptions, or even theories, bristling with important, but perhaps hidden, implications. It is these implications which effectively force the parties in the original position to disregard the form of the basic structure, leaving it to be defined in accordance with historical dominance.

That the circumstances of justice are assumptions is not a problem in itself; every argument requires premises. It is Rawls's conception of the nature of his enterprise that creates the difficulties to be spelt out below. He intends that the argument for justice as fairness should be “strictly deductive” and cites theories in politics and economics(!) as evidence of the fact that fully deductive theoretical schemes can be based on assumptions about human beliefs and interests. His ultimate aim is to achieve “a kind of moral geometry” with “all the rigor which this name connotes” (TJ p121).

The problem of course is that the premises of Rawls's argument are poor analogues for the axioms of geometry. The latter are simple inventions which carry little,
if any, implicational baggage, and the primitives involved (such as 'point' and 'line') need not necessarily map onto features of observable reality as closely as the parties in the original position are meant to be connected to actual persons. By contrast, it will be argued that the circumstances of justice smuggle tacit premises into the deductive scheme, premises that skew the argument towards the position where the question of whether the just society is to be capitalist, socialist, or any other kind, is of little interest or import. The SCJ imply an unacknowledged restriction on the ends of the parties in the original position, in defiance of Rawls's self-imposed methodological precepts. And the OCJ unjustifiably render the question of property rights in the means of production superfluous in the deliberative stage via undefended assumptions about the nature of technology and its relationship to the institutions of production. This contention, with regard to the SCJ, will now be elucidated.

Outcomes, Forms and Preferences

It is significant that Rawls stresses the 'competing claims' aspect of the SCJ. This, in itself, suggests that the parties in the deliberative stage of the original position will have an overriding interest in the distribution of primary goods. This impression is confirmed by Rawls's account of the reasoning behind the two principles of justice, especially that of the difference principle.

According to Rawls the benchmark position in the deliberative stage is one of complete equality in the distribution of primary goods (TJ p62, PL p282). From this position, departures from equality are considered, both in "organisational powers" - an undefined term - and the distribution of wealth and income (TJ p62). Where these inequalities work to the benefit of the least-advantaged representative person, they ought to be permitted. The extent of inequality permitted is determined by the point at which greater inequality fails to improve the lot of the least-advantaged.

---

8 Euclidean geometry is a pure axiomatic system; replace the geometrical terms consistently in the postulates with nonsense syllables and the system still 'works' in that all the conclusions are logically valid. If 'persons' is replaced by booyangs' throughout TJ or PL, the result is morass of pointless verbiage. It is the fact that his argument is intended to be in some way descriptive, and applicable to the world, which makes Rawls's endeavour something different in kind from pure axiomatic systems like geometry.

9 An interesting case concerns the objective circumstances, and the assumption of moderate scarcity. On what basis is this known to be true? Reflection on the history of human societies seems to imply that an imbalance between demands and the extent of the social product is the norm, but is this to be taken therefore as an enduring 'fact' about all societies past, present and future? Why must it be true that scarcity will prevail? This 'fact' is in reality an inductive argument with no particular claim to universal truth. This observation does not commit us to the Utopian proposition that scarcity definitely can be eliminated, but rather it points to the fact that Rawls bases his "moral geometry" not on facts per se, but, in this case, on an argument taken at face value.

10 The case of the OCJ is taken up in detail by Bradford (2000a).

11 This is especially evident when he endeavours to encapsulate the meaning of the circumstances of justice. Thus, "one can say, in brief, that the circumstances of justice obtain whenever mutually disinterested persons put forward conflicting claims to the division of social advantages under conditions of moderate scarcity" (TJ p128), and, "justice is the virtue of practices where there are competing interests and where people feel entitled to press their rights on each other" (TJ p129).

12 Rawls speaks of the basic structure distributing primary goods, but it is clear that this is true only of a subset (TJ p62). Rights and liberties are part of the 'skeleton' of the basic structure, being guaranteed to all via the constitution and just legislation. Thus the distributive aspect of the basic structure comprises the mechanics of wealth and income distribution, and the means whereby powers and opportunities (the worth of liberty) are shared out. Clearly the former is the most important with respect to discussions of the difference principle, and therefore to the central theme of this paper.

13 Rawls ties the welfare of the least-advantaged to that of all other groups in society by means of the
When turning to a more detailed discussion of this point, Rawls acknowledges that the meaning of a departure from distributional equality being to 'everyone's advantage' is open to several interpretations (TJ p65). He therefore considers a variety of interpretations of the Second Principle (TJ pp. 65-83). The common thread running through all of these is what Rawls calls the “principle of efficiency” (TJ pp. 67-68). This is none other than the familiar concept of Pareto optimality, suitably modified to apply to the basic structure rather than the allocation of a stock of commodities among a group of individuals.

The system of Natural Liberty is the libertarian model of society. With careers open to those with the talent and desire to strive for them, the distribution of wealth, income and authority is determined by the principle of efficiency, in the context of a free market system, and the result deemed just. The system of Liberal Equality is an attempt to counter the moral arbitrariness of Natural Liberty by insisting on fair, and not simply formal, equality of opportunity in the pursuit of social positions. In this way the contingencies of social fortune are countered, and the resulting Pareto optimal distribution reflects the underlying distribution of natural abilities.

The system of Democratic Equality combines fair equality of opportunity with the difference principle. In effect, the just distribution is that Pareto optimum which satisfies first the condition that it has been achieved through fair equality of opportunity, and second that the expectations of the least-advantaged representative person are maximised. Rawls obviously concludes that this interpretation is to be preferred and presented to the parties in the deliberative stage.

Thus, in line with the 'competing claims' aspect of the SCJ, the reasoning for the principles of justice is based squarely on consideration of the distribution of primary goods, and not at all on the manner in which they are to be produced. The underlying reason for this is straightforward: Rawls assumes that the parties in the deliberative stage accept the Pareto principle as the yardstick by which to judge the efficiency of social and economic arrangements (TJ p67).

Pagano (1985) demonstrates convincingly that modern general equilibrium theory, and thus its associated welfare economics, relies on the assumption that work assumption of chain connection (TJ pp. 81-82). This principle entails that, at least over a certain range, the welfare of the various socioeconomic ‘groups’ (from 'best-off' to 'worst-off') does not display mutual independence. Specifically, anything that improves the situation of the worst-off can be relied upon to do the same for the better-off, and vice versa. As a result, departures from the complete-equality benchmark can be considered either to everyone’s advantage or not, depending upon their effect on the least-advantaged. Barry (1973 pp. 110-11) labels chain connection “a sort of living fossil, a coelacanth among ideas” and notes that it

\[\text{instantly knocks the bottom out of most possible issues concerning the distribution of desirable but scarce resources of the kind which have given an edge to political dispute in the past few millennia.}\]

Rawls explicitly declines to examine the likelihood that chain connection will, or does, hold (ibid.).

\[\text{Rawls acknowledges the index-number problem associated with assessing the welfare of the least-advantaged based on different 'baskets' of primary goods (TJ pp. 93-5). What he does not comment on, however, is exactly how the parties in the deliberative stage know that an unequal distribution impacts upon the welfare of the least-advantaged. How is it that they are aware of the mechanism through which more for some makes others better off? They have neither information about the nature of their society nor any about themselves or any other individual. They do have access to the principles of social, political and economic theory, but it is hard to see how the required inferences could be made in the absence of more detailed empirical information.}\]

\[\text{These are expectations of realising their conception of the good, which are dependent on the bundle of primary goods received.}\]
enters the preference structure of economic agents only in regard to the amount of labour supplied by the individual, and not in relation to the nature of the tasks performed. This aspect of orthodox theory is rooted in Walras's treatment of labour whereby the labour supply of an individual is divided into two parts. One part is sold for use in the production of consumer goods or consumption services, while the remaining part is 'consumed' by the individual him(her)self. This latter part has subsequently been labelled 'leisure' (or some utility-producing 'reservation activity') and is assumed to yield utility to the individual, while that part of available labour which is dedicated to the production of consumer goods is assumed to have no impact on the welfare of the individual concerned.  

Thus the disutility of labour supply consists only of the opportunity cost of the number of hours worked in terms of foregone leisure. The use of the 'leisure device' implies that workers are indifferent to their conditions of work and the nature of the tasks they are to undertake; 8 hours working in a sewage treatment plant is therefore a no less desirable prospect than the equivalent amount of time spent working as a photographer of supermodels. Consumption goods and leisure time are the only ends desired by these agents, and work is regarded purely as a homogeneous means to these ends.

Pagano demonstrates these implications more formally in the context of an orthodox resource allocation model. He shows that when the social welfare function is dependent on the output vector of goods only, where leisure is a good produced by a reservation activity, technical efficiency is a necessary condition for social welfare maximisation. The optimisation problem can be broken down into two independent steps. First, purely as an engineering problem, the set of efficient vectors is identified; the quantity of manpower available to each firm is allocated among productive uses in such a way that more of one commodity can only be produced if the output of at least one other is reduced. Once this is achieved, preferences over output are considered in order to identify the efficient vector which maximises social welfare.

In other words, as preferences over the organisation of production - that is, the allocation of tasks among jobs - play no part in the determination of the efficient production vectors, social welfare is invariant to the allocation of manpower within firms. This result holds both for capitalist firms and the single-firm centrally-planned economy. Under both regimes there is the possibility of rational conflict over the distribution of output, but none whatever over the organisation of production as the latter affects no individual's welfare.

Now let us return to Rawls and the SCJ. If attention is focussed on the 'economic' primary goods (wealth and income) it is clear that the 'competing claims' aspect of the SCJ emphasised by Rawls is just another means of expressing the orthodox economic

---

16. For example, Koopmans (1957 p73n), Rawls's main reference for economic theory, acknowledges the need to introduce "artificial commodity concepts": "if labor is a primary input but leisure a desired commodity, one may introduce 'man-hours' as a primary commodity and have its use as labor in actual production compete with its use in a 'process' which converts 'man-hours' into an equal number of desired 'leisure-hours'".

17. In Pagano's model the production function is defined over the vector of tasks. A task is understood to be a "use of manpower in production" described by a specification of the work to be done, how it is to done, and the time permitted for its execution (1985 p118). From the definition of tasks are derived the definitions of a job and a skill. The former is a subset of tasks to be performed by one unit of manpower within a particular firm, and thus represents the organisation of production within the firm. A skill is the resource endowment of an individual worker, defined as the set of tasks that the individual unit of manpower is able to perform.
assumption that preferences over work do not enter the social welfare function, that
distribution, not production, is the prime concern.\footnote{18}

This flies in the face of Rawls's insistence that he would not impose constraints
on the conceptions of the good that are held (though unknown) by the parties in the
deliberative stage. Apparently they are not to include the characteristics of the work they
perform among the things they will value beyond the veil of ignorance. Furthermore,
given that the parties must possess some kind of general awareness of the nature of
human life and experience, the assumption that the quality and enjoyability of work is to
be accorded no importance whatever is a strong and startling one to make.\footnote{19}
This is in
direct conflict with Rawls's claim that the case for justice as fairness is based on “widely
shared and yet weak conditions” and assumes “as little as possible” (\textit{TJ} p129).
Rawls
appears to have flouted his own methodological prescriptions.

This inconsistency notwithstanding, Rawls faces more fundamental difficulties
arising out of his version of the SCJ. Pagano (\textit{op.cit.}) reveals that implicit in resource
allocation models which exclude preferences over work is the principle that de-skilling
of jobs is conducive to the achievement of higher levels of social welfare.\footnote{20}
Within the
framework of Pagano's model, de-skilling entails the redefinition of jobs within the
production process of a firm (set of tasks to be performed) such that the number of tasks
comprising at least one job is reduced, without an increase in the tasks comprising any of

\footnote{18} It is fair to assume that Rawls is aware of this assumption and its role in orthodox theory, as it is
mentioned at least twice in the section from Koopmans to which Rawls refers on the question of the
principle of efficiency. There Koopmans notes that the optimality results of neoclassical theory require
that “the satisfaction of each consumer is affected only by his own consumption and work, and not in
addition by other people's consumption or by any production processes except through the fact that he
contributes labor” (Koopmans 1957 p41, emphasis added). Later Koopmans reiterates that the proof of
the allocative efficiency of a competitive market organisation depends in part on the assumption of the
“independence of any man's [sic] preference structure from any production process”(p53).

\footnote{19} Given the history of industrial relations in the United States, marked as it is by outbreaks of violent
and often fatal conflict over the organisation of production, it is a patently absurd assumption for Rawls
to make. The historical record on this issue also poses some interesting questions with respect to the
constitutional and legislative stages. When historical information is made available in these stages,
facts of this nature, common to most industrialised countries, will presumably become known. Would
we not expect this to influence any decisions regarding the ownership of the means of production? If,
historically, capitalism has been characterised by violent struggles between the owners of the means of
production and those whose labour they purchase, is that not a good reason to consider alternative
property rights arrangements? Rawls's reliance on historical dominance implies that historical forms
will be perpetuated despite past failings; at no time does he explicitly acknowledge the possibility that
a wholesale change of regime may be required by the principles of justice.

A more important question involves the extent to which information about individual
conceptions of the good may be inferred from the historical information available in the later stages of
the original position. If the parties become aware that many individuals have risked death and injury in
the name of workplace reform and workers' rights, is it not reasonable to assume that they will infer,
correctly, that the quality of work is an important element of the conception of the good of at least
\textit{some} persons? Once this step is taken, however, it is hard to see how the issue of the form of the basic
structure would not become extremely important in the original position. More fundamentally still,
once the parties have deduced that the organisation of production features in conceptions of the good,
would they not consider Rawls's account of the SCJ to be incomplete and inaccurate? As the SCJ are
one of the empirical constraints on the deliberative stage, this situation would necessitate a revaluation
of the principles of justice in light of the new information.

\footnote{20} As Pagano explains, this idea is rooted in pre-neoclassical thought. Gioia argued that a detailed
division of labour would yield savings on training time, and thus be of benefit to society as a whole.
Babbage maintained that the 'master manufacturer' could, assuming wages proportional to the skill
requirements of jobs, realise cost savings by breaking complex and demanding tasks down into skilled
and unskilled components. James Mill held that the Babbage principle should be extended to society as
a whole in order to enhance social welfare through increased production of consumption goods.
the remaining jobs. As de-skilling enlarges the set of employable skills in an economy, it eases skill constraints and permits an increase in output from given manpower inputs. Given a social preference structure defined over goods only, this implies that de-skilling can only increase social welfare, or at least leave it unchanged (Pagano 1985 pp. 122-126). Since the set of feasible output vectors for an individual firm remains feasible after de-skilling, while new vectors may be attainable, the process of de-skilling cannot decrease profits. Pagano states that

the conclusion that can be derived from the resource allocation model is that the internal allocation of manpower will be organised by a profit-maximising manager according to the criterion that the de-skilling of jobs is never disadvantageous and that such an organisation of manpower is consistent with the aim of improving social welfare (ibid.).

Moving beyond pure economic theory, these conclusions are also supported strongly by the labour process theory literature, particularly Braverman (1974). Braverman focuses on the implications of Taylorism, or 'scientific management' as an extension of the Gioia-Babbage principle. He isolates three main principles of Taylorism - the dissociation of the labour process from the skills of the workers, the separation of the conception and execution of tasks, and the use of management's knowledge of the productive process to plan and control production in a stepwise process - and argues that their widespread implementation in the capitalist mode of production reflects the necessity of control over the labour process under capitalism. Control is necessary if capitalists are to realise the potential of purchased labour power, impose their will within the social relations of production, and lower production costs (increase profits) by rendering labour more homogeneous. He concludes that the proliferation of Tayloristic control mechanisms has resulted in a destruction of all-round skills and the degradation of work.

Braverman updates Marx's analysis of the labour process by examining the degradation of work in the history of twentieth-century industrialisation. As much of the labour process literature is also based extensively on historical information - for example, Marglin (1974) - the conclusions therein would presumably become known in the legislative stage, when the decision regarding the ownership of the means of production is meant to be made. Since the conclusions reached are scarcely flattering to capitalism, it would seem that Rawls is required to explain why historical dominance holds in the face of such information. If historical material of this nature is to be excluded from the legislative stage, Rawls must explain the criteria that differentiate 'acceptable' facts about the past from those that are 'unacceptable'. Without such explanations, Rawls's conclusions seem to be ultimately little more than an elaborate expression of his personal beliefs.

Thompson (1989) notes that many within the literature on labour process theory have criticised what they see as the inexorable trend toward de-skilling implied by Braverman. He acknowledges the importance of national differences in the development of the labour process (particularly with respect to the Japanese experience), the impact of flexible specialisation and the associated trend toward multis killing and the recombination of intellectual and manual skills, and the incidence of both worker resistance to control and cooperation in the re-shaping of the labour process.

---

21 For example, see the papers collected in Wood (1982).
He concludes, however, by noting that the rejection of a uni-dimensional version of the Braverman thesis does not thereby negate the important conclusions of the labour process literature. When, as under capitalism, the capacity to work is utilised as a means of producing value, the logic of accumulation forces a constant revision of the labour process in order to accommodate the competition between units of capital, and the antagonism between capital and labour that characterises the capitalist mode of production. Linked to this is a control imperative; structures of control over labour are necessary if legal ownership of the means of production is to be translated into real possession. Neither of these principles imply that simple de-skilling is their only possible manifestation; capitalists will engage in flexible strategies which can incorporate re-skilling and the ceding of selected areas of responsibility and control. Overall, however, the need to retain general directive control over labour, and to satisfy the demands of accumulation, implies that both the use of workers' individual creative capacities, and attempts to dismantle hierarchical relations, must play a limited role in the capitalist labour process.

The problem that the theoretical results of Pagano and labour process theory pose for Rawls concerns his account of the rationality of the parties in the original position, the so-called 'thin theory of the good'. Rawls considers that a person's good is determined by the rational plan of life that he/she would choose with deliberative rationality from the maximal class of plans (TJ p424). A life-plan is seen as a hierarchy of sub-plans reflecting a procession of desires from more to less general. Although plans cover the whole of an individual's life, the details of the various sub-plans decrease the further they are away from the present. Rational plans are those which most efficiently match means and ends, and encompass the results of other plans (Rawls calls the latter the principle of inclusiveness). Rawls insists that the theory of the good must be 'thin', meaning that ideas of human excellences, and other specifications of what ought to be wanted, should not be incorporated into the definition of rationality.

Yet without some restriction on the content of rational plans, Rawls's argument is in danger of collapsing into a kind of utilitarianism; the good for an individual simply becomes getting more of what he/she wants (Barry 1973 p30). In such a situation, there is no reason for the parties in the deliberative stage to focus on the primary goods as the things to be distributed. As a result, the solution to their choice problem (deciding on

---

22 The concept of deliberative rationality, which Rawls ascribes to Sidgwick, is in essence a state of perfect foresight with respect to the outcomes of plans. The problems associated with defining 'goodness' for parties in the original position with reference to the case where all future outcomes are known need little elaboration. The parallel with the role of perfect foresight assumptions in (much) neoclassical theory is intriguing.

23 There are interesting similarities between Rawls's device of rational plans and Koopmans's discussion of the interpretation of the competitive general equilibrium model with respect to time. Rawls has the conception of the good determined by the rational life plan that he/she would choose with deliberative rationality from the maximal class of plans (TJ p424). A life-plan is seen as a hierarchy of sub-plans reflecting a procession of desires from more to less general. Although plans cover the whole of an individual's life, the details of the various sub-plans decrease the further they are away from the present. Koopmans notes that when time is divided into successive periods and commodities dated, the choice of a consumer is in fact "a plan for future consumption extending over all periods considered". Furthermore "[the consumer's] preference ordering is thought of as an ordering of all such plans". Koopmans (1957 p61) admits that uncertainty is excluded from the model: "[c]onsumers choose a plan for their lifetime, in full present knowledge of their future preferences, of the time of their death, and of the place occupied in their preference structure by the resources handed on at death to their heirs". Rawls does not acknowledge any debt to Koopmans, nor is any suggested, but it is intriguing to consider that the thin theory of the good may be based in part on some aspects of intertemporal general equilibrium theory.
principles of justice) is indeterminate; the argument for justice as fairness stalls at the outset.

Since the parties in the deliberative stage are assumed not to know the details of their conception of the good, Rawls is forced to introduce an empirical generalisation about human psychology that is to be known in the original position, and which will imply that rationality entails desiring the primary goods. This is the Aristotelian Principle: *ceteris paribus*, humans enjoy the exercise of their innate and trained capacities, and this enjoyment is proportional to the extent of the capacity utilised and the degree of complexity required (TJ p414). For example, a person who can play both draughts and chess would prefer the latter. From this Rawls concludes that humans prefer a larger pattern of ends (in order to utilise capacities more intensively) and that plans will reflect this in that comprehensive plans will be preferred.\(^{24}\) In other words, humans have a higher-order desire to follow the principle of inclusiveness.

For Rawls, the Aristotelian Principle is a theory of motivation that accounts for our major desires and preferences (TJ p428). By ordering all human activities in interdependent chains according to the principle of inclusiveness - such that the \(n^{th}\) activity in a given chain encompasses all of the skills required for the \((n-1)^{th}\) - the Aristotelian Principle can be reformulated as the assertion that humans always desire to move up chains. In fact, in a manner sure to please any neo-classical economist, humans will apply their available resources (in terms of time and energy devoted to development) to moving up the chains associated with any preferred pattern of activities to the point where no benefit is to gained from rearrangement (TJ p431).

The role of the Aristotelian Principle in the thin theory of the good is, according to Rawls, that it

states a deep psychological fact which, in conjunction with other general facts and the conception of a rational plan, accounts for our considered judgements of value. The things that are commonly thought of as human goods should turn out to be the ends and activities that have a major place in rational plans. The principle is part of the background that regulates these judgements (TJ p432).

The question of the relationship between the thin theory of the good, centred on the Aristotelian Principle, and the rationality of desiring primary goods is unclear.\(^{25}\) In

\(^{24}\) Note the difficulty associated with the fact that it is the *content* of rational plans (or specifically the plan which would be chosen with deliberative rationality) which defines the good for a person. What then is the basis for preference between plans? How can we make sense of the claim that comprehensive plans are 'better' than simple plans? It cannot be that they produce more 'goodness', as that would be circular.

\(^{25}\) Rawls himself seems undecided of the truth of the Principle and indeed whether or not it needs to be true in the context of his theory. On page 427 he claims that there is no reason to explain why it is true. After introducing some casual, anecdotal 'evidence', he concludes that the issue is unimportant for the theory of the good. On page 429 the Principle has become a tendency rather than an invariable pattern, one which could be overridden by the costs of training. Two pages later Rawls alludes to evolution to support the Principle, and holds that the important point is to illustrate its role in the formulation of rational plans *given that it is true*. On page 432 he makes the truth of the Principle a condition for it having a proper role in moral theory, despite having just declared that it expresses a "deep psychological fact". He then goes on to state that even if the Principle is not universally true, the idea of rational plans as the determinants of goodness is still valid. He strengthens this claim on the next page by holding that the correctness of defining goodness in terms of rational plans does not require the truth of the Aristotelian Principle at all. In fact it is defensible on instrumental grounds as by assuming it we can explain why certain things have been recognised as goods. Moreover, it has a central role to play in the moral psychology upon which the theory of justice as fairness is based. Economists will be reminded at this point of Hicks's (1946 p84) admission that the main reason for maintaining the
fact, Rawls declines to make the case, claiming that the relationship is “evident enough” (TJ p434). The impression given by the arguments for goodness as rationality is that the primary goods are means to virtually all ends, and since the Aristotelian Principle implies that rational plans will be those which include the widest possible diversity of ends, it follows that rational parties in the original position would wish to secure as large an amount of primary goods for themselves as possible.

Thus, as a motivational assumption, the Aristotelian Principle explains the desire for primary goods in the deliberative stage. Unfortunately, it conflicts with the implications of the economic and social theory presumed known in that stage. Given that they know humans enjoy exercising their abilities, how will the parties react to the fact that the logic of capitalist production requires the de-skilling of jobs, and thus the retardation of those skills, the work-related skills, utilised most intensively throughout a human life? If the Aristotelian Principle is indeed a theory of motivation that accounts for our major desires and preferences (TJ p428), why would the implications of capitalist ownership and control mechanisms, with regard to de-skilling and the maintenance of authoritarian and hierarchical relations within production, not result in the question of the organisation of production taking a prominent place in the discussions of the deliberative stage?

This conclusion is strengthened by the psychological evidence, cited by Bowles and Gintis (1976), Ryan (1981) and Schwartz (1982), to the effect that participation in monotonous, routine activity is damaging to cognitive performance in general. Thus the process of de-skilling could be held to lead to a general deterioration in the ability of workers to exercise their capacities in any and all of the activities in which they engage. Awareness of this potential effect of capitalist employer behaviour represents an avenue through which the assignment of property rights over the means of production could affect the development of the “sufficient material basis for a sense of personal independence and self-respect” that Rawls holds to be essential for the development of the two moral powers (PL p298). Possession of the two moral powers is held by Rawls to be a necessary and sufficient condition for being regarded as a “full and equal member of society” with respect to questions of political justice, and thus is central to the development of self-respect, the most important of primary goods (PL p302). As such, the definition of rationality for the parties in the
deliberative stage would encompass an interest in the ownership status of the means of production.

**Difficulties and Remedies**

If this were to occur, however, Rawls’s argument comes apart in a number of crucial places. First, it is unclear how the central focus on the primary goods could be justified in this context. Second, the difference principle could no longer be defined in relation to the principle of efficiency; the parties in the deliberative stage would no longer be satisfied merely with considering deviations from a benchmark of an equal *distribution* of primary goods, as the form of the basic structure would be an important, relevant consideration. A new method of justification is required. Third, the problems associated with the specification of the basic liberties would be intensified, as there seems to be no obvious reason to exclude social property in the means of production from the list. Yet allowing its inclusion implies that it should be incorporated into the just constitution. This would have the effect of negating Rawls’s conceptual division of the basic structure into ‘political’ and ‘economic’ sections, and, more fundamentally, would undermine the priority rules of justice as fairness whereby political justice is lexically prior to equality of opportunity which is lexically prior to the justice of economic distribution.

In light of these disastrous consequences, how can Rawls respond? Unfortunately, given the structure of the argument in *TJ*, his options are limited. Altering his account of the SCJ so as to take account of preferences over work is not a feasible option. Pagano (1985 p126-29) shows that when workers’ preferences over tasks are included in the social welfare function, the manpower allocations of profit-maximising firms need not be conducive to social welfare maximisation. As de-skilling can now result in a *decrease* in social welfare, the derivation of the optimum output vector can no longer be decomposed into a two-stage procedure beginning with the isolation of efficient vectors, conceived as an engineering problem. Hence, Rawls could no longer employ the principle of efficiency in deriving the difference principle and he would have to recast his entire argument, taking into account the subsequent interest the parties in the deliberative stage would have in the organisation of production. The dilemma facing Rawls may be put as follows. If he maintains his account of the SCJ and ignores preferences over work, then the conflict between the Aristotelian Principle and the de-skilling implications of economic theory renders untenable arguments based on the principle of efficiency, as the parties will not accept simple comparisons of distributions only. If he repudiates his version of the SCJ and recognises the welfare implications of work, the principle of efficiency ceases to be operative and thus cannot be utilised in deriving the principles of justice.

As the inconsistency between the SCJ and the Aristotelian Principle is at the root of Rawls’s difficulties, he could utilise another principle of motivation. Although he acknowledges in *PL* (p203n) that other, unspecified, psychological principles could yield the same results, it is not really possible to deduce the ramifications of such a move for Rawls’s system. It is worth noting, however, that in *PL* (pp. 203, 207) Rawls indicates

---

29 Throught (1980-81) employs the concept of “psychic income” to derive the same conclusion and argue that capitalist work organisation represents a case of market failure. In a perspicacious treatment of the issue, Sen (1975) demonstrates the problems posed for assessments of ‘efficiency’ by the existence of worker preferences relating to both tasks and alternative techniques. Albert and Harhnel (1991 p19) argue that tasks that are “employee empowering” in their effects will be undersupplied, and paid less than the socially optimal wage, under capitalism, thus generating a form of market failure.
that the Aristotelian Principle is central to the case for the goodness of the well-ordered society of justice as fairness. The link is through the exercise of the two moral powers, which are interpreted as examples of the kind of capacities to which the Aristotelian Principle refers. Hence certain key aspects of Rawls's program seem to be dependent on the particular moral psychology represented by the Aristotelian Principle.

Buchanan (1982 pp. 123-124) suggests that Rawls could evade the difficulties outlined above, with respect to the SCJ, by including 'meaningful' work as one of the social bases of self-respect, the most important of the primary goods. However, this defence is unreliable for several reasons. First, Rawls holds that the definition of meaningful work is not a problem of justice (TJ p290). As such, there seems to be little scope for incorporating it into the list of primary goods. In any case, Buchanan offers no account of how the thin theory of the good, as it stands, provides a justification for considering 'meaningful' work to be a primary good. Nor does he explain how the reasoning for the principles of justice could incorporate the link to the organisation of production implicit in a definition of rationality that refers in part to a desire for 'meaningful' work. In short, Buchanan may wish to espouse a mutant form of Rawls's argument, but Rawls cannot.

Second, although Rawls identifies 'meaningful' work as one of the 'human goods' (TJ p425), it is clear from the context that it is an end rather than a means. As such, including it among the primary goods serves only to make the thin theory of the good appreciably 'thicker', a state of affairs precluded by Rawls's methodology.

Third, it seems a rather ad hoc procedure to simply enlarge the list of primary goods whenever a critic points out lacunae in the argument. If the concept of primary goods is to be useful, then surely the list of primary goods must be a limited subset of all possible goods. If the two are coextensive, then justice as fairness becomes utilitarianism. If most goods are considered primary then the coherence of the thin theory of the good is endangered. Rawls himself is guilty of the sin of multiplying the primary goods without necessity. When replying to Musgrave (1974) on a point concerning optimal taxation, defensive goods-leisure substitutions and the institutions of the just society, Rawls (1974 p654) indicates that leisure may be included in the list of primary goods. The problem with this case is that the basic structure is held by Rawls to distribute the primary goods. If leisure is construed as a strictly private activity, as in mainstream economic theory, then it is not distributed by the major social and economic institutions, or anything else. However, not regarding leisure as a private activity in that sense would require attention to the organisation of production, as it is the economic institutions that will 'distribute' leisure through their impact on required work time.

Finally, the link to self-respect is not as obvious as it seems on the basis of 'commonsense' psychology. For Rawls, self-respect is derived not from interpersonal experiences directly, but from knowledge of the justice of the basic structure. Hence, by securing the basic rights and liberties for all, and guaranteeing fair equality of opportunity, the just society underwrites the “essentials of persons' public recognition as free and equal citizens” (PL p203). This state of affairs secures the fundamental needs of citizens and establishes the social bases of self-respect. Clearly, there are few prospects for integrating 'meaningful' work into such a conception of self-respect.

In presenting the next possible line of defence, it should be noted that Rawls is aware of the less desirable implications of the capitalist labour process, and, indeed, he feels they would be inappropriate in the just society. He declares that:

---

30 Rawls states elsewhere that “what men want is meaningful work in free association with others, these associations regulating their relations to one another within a framework of just institutions” (TJ p290).
31 Private in the sense that it involves a given, private 'reservation activity'.
a well-ordered society does not do away with the division of labor in the most
general sense. To be sure, the worst aspects of this division can be surmounted:
no one need be servilely dependent on others and made to choose between
monotonous and routine occupations which are deadening to human thought and
sensibility. Each can be offered a variety of tasks so that the different elements of
his nature find a suitable expression (TJ p529).

The sentiment is welcome, but Rawls provides no explanation as to how the
monotonous jobs will be replaced by stimulating and rewarding ones in the just society.
Barry (1973 pp. 163-165) is probably correct when he suggests that Rawls believes that
ensuring the freedom of choice of occupation, and full employment through the
stabilisation branch, will suffice to make work meaningful. If this were so, then Rawls
could evade the criticisms made above. He certainly argues that, in equilibrium, the
attractiveness of jobs will be equalised (TJ p305). This ties in with his partiality to the
property-owning democracy of Meade (1964). Meade waxes utopian about the way in
which work would become purely voluntary in such a system (exit would be costless).32

But since the characteristics of jobs are out of the control of workers it seems
there is no mechanism to ensure that jobs are meaningful.33 Such an outcome may be
possible if jobs covering all conceivable combinations of tasks were available under
conditions of continuous full employment. Given the difficulty of achieving the latter
conditions under capitalism, and the fact that the imperatives of control and
accumulation make it unlikely that all superseded production methods would coexist
with their replacements, this seems a most unhelpful straw at which to clutch.

Rawls's final, and most feeble, retort is methodological in character. He argues
that common understanding in the original position requires bounds on the complexity of
theoretical arguments permitted there (TJ p142). Although it is difficult to grade general
facts on the basis of complexity, he holds that we “recognise an intricate theoretical
construction” easily enough. On these grounds, he supposes that conceptions of justice
which depend on simpler general facts, and which require less trawling through
theoretically defined possibilities, are to be preferred. That the author of the 582 pages of
argument which constitute TJ could put forward this proposition without apparent irony
is, to put it mildly, quite surprising, but in any case it is somewhat hypocritical to reject
the above criticisms on the ground that they require excessive theoretical manipulation.
At root, the criticisms rest on aspects of the very theories that Rawls calls upon to justify
his own argument. Besides, whether or not the economics of Koopmans is considered an
‘intricate theoretical construction’, it is certainly no more so than the argument for justice
as fairness itself.

Conclusion

Ultimately, Rawls is left with no viable defence to the charge that the neglect of
production in favour of distribution in his system stems from his 'competing claims'
account of the SCJ, which in turn reflects the assumption of neo-classical economic
theory that social welfare and technical efficiency are independent of preferences over
work. The implications of the SCJ clash with the psychological premises upon which the
rationality of parties in the original position is based, and thus deprive Rawls's

---
32 Bradford (2000b) examines the relationship between Rawls and Meade in greater detail.
33 Perhaps the orthodox model assumes that jobs, like consumption goods, have only one characteristic
   - in this case 'jobness'?
conclusions of logical support. One or the other must go, but modifying the SCJ (retreating from the 'competing claims' interpretation and taking account of preferences over the organisation of production) means ditching the principle of efficiency and therefore the argument for the principles of justice, particularly the difference principle.

Thus Rawls appears committed to retaining his account of the SCJ if he is to preserve, at least superficially, the integrity of his argument. This leaves him in the awkward position of maintaining that the parties in the original position will focus on distribution, and ignore the form of the basic structure, only on the basis of a strong assumption that represents a prior constraint on possible conceptions of the good in violation of his own methodological dicta.
References


