

# CHAPTER

## 7

---

# Distributive Justice

---

**T**HE minimal state is the most extensive state that can be justified. Any state more extensive violates people's rights. Yet many persons have put forth reasons purporting to justify a more extensive state. It is impossible within the compass of this book to examine all the reasons that have been put forth. Therefore, I shall focus upon those generally acknowledged to be most weighty and influential, to see precisely wherein they fail. In this chapter we consider the claim that a more extensive state is justified, because necessary (or the best instrument) to achieve distributive justice; in the next chapter we shall take up diverse other claims.

The term "distributive justice" is not a neutral one. Hearing the term "distribution," most people presume that some thing or mechanism uses some principle or criterion to give out a supply of things. Into this process of distributing shares some error may have crept. So it is an open question, at least, whether *redistribution* should take place; whether we should do again what has already been done once, though poorly. However, we are not in the position of children who have been given portions of pie by someone who now makes last minute adjustments to rectify careless cutting. There is no *central* distribution, no person or group entitled to control all the resources, jointly deciding how they are to be doled out. What each person gets, he gets from others who give to him in exchange for something, or as a gift. In a free soci-

ety, diverse persons control different resources, and new holdings arise out of the voluntary exchanges and actions of persons. There is no more a distributing or distribution of shares than there is a distributing of mates in a society in which persons choose whom they shall marry. The total result is the product of many individual decisions which the different individuals involved are entitled to make. Some uses of the term "distribution," it is true, do not imply a previous distributing appropriately judged by some criterion (for example, "probability distribution"); nevertheless, despite the title of this chapter, it would be best to use a terminology that clearly is neutral. We shall speak of people's holdings; a principle of justice in holdings describes (part of) what justice tells us (requires) about holdings. I shall state first what I take to be the correct view about justice in holdings, and then turn to the discussion of alternate views.<sup>1</sup>

## SECTION I

### THE ENTITLEMENT THEORY

The subject of justice in holdings consists of three major topics. The first is the *original acquisition of holdings*, the appropriation of unheld things. This includes the issues of how unheld things may come to be held, the process, or processes, by which unheld things may come to be held, the things that may come to be held by these processes, the extent of what comes to be held by a particular process, and so on. We shall refer to the complicated truth about this topic, which we shall not formulate here, as the principle of justice in acquisition. The second topic concerns the *transfer of holdings* from one person to another. By what processes may a person transfer holdings to another? How may a person acquire a holding from another who holds it? Under this topic come general descriptions of voluntary exchange, and gift and (on the other hand) fraud, as well as reference to particular conventional details fixed upon in a given society. The complicated truth about this subject (with placeholders for conventional details) we shall call the principle of justice in transfer. (And we shall suppose it also in-

cludes principles governing how a person may divest himself of a holding, passing it into an unheld state.)

If the world were wholly just, the following inductive definition would exhaustively cover the subject of justice in holdings.

1. A person who acquires a holding in accordance with the principle of justice in acquisition is entitled to that holding.
2. A person who acquires a holding in accordance with the principle of justice in transfer, from someone else entitled to the holding, is entitled to the holding.
3. No one is entitled to a holding except by (repeated) applications of 1 and 2.

The complete principle of distributive justice would say simply that a distribution is just if everyone is entitled to the holdings they possess under the distribution.

A distribution is just if it arises from another just distribution by legitimate means. The legitimate means of moving from one distribution to another are specified by the principle of justice in transfer. The legitimate first "moves" are specified by the principle of justice in acquisition.\* Whatever arises from a just situation by just steps is itself just. The means of change specified by the principle of justice in transfer preserve justice. As correct rules of inference are truth-preserving, and any conclusion deduced via repeated application of such rules from only true premisses is itself true, so the means of transition from one situation to another specified by the principle of justice in transfer are justice-preserving, and any situation actually arising from repeated transitions in accordance with the principle from a just situation is itself just. The parallel between justice-preserving transformations and truth-preserving transformations illuminates where it fails as well as where it holds. That a conclusion could have been deduced by truth-preserving means from premisses that are true suffices to show its truth. That from a just situation a situation *could* have arisen via justice-preserving means does *not* suffice to show its justice. The fact that a thief's victims voluntarily *could* have presented him with gifts

---

\* Applications of the principle of justice in acquisition may also occur as part of the move from one distribution to another. You may find an unheld thing now and appropriate it. Acquisitions also are to be understood as included when, to simplify, I speak only of transitions by transfers.

does not entitle the thief to his ill-gotten gains. Justice in holdings is historical; it depends upon what actually has happened. We shall return to this point later.

Not all actual situations are generated in accordance with the two principles of justice in holdings: the principle of justice in acquisition and the principle of justice in transfer. Some people steal from others, or defraud them, or enslave them, seizing their product and preventing them from living as they choose, or forcibly exclude others from competing in exchanges. None of these are permissible modes of transition from one situation to another. And some persons acquire holdings by means not sanctioned by the principle of justice in acquisition. The existence of past injustice (previous violations of the first two principles of justice in holdings) raises the third major topic under justice in holdings: the rectification of injustice in holdings. If past injustice has shaped present holdings in various ways, some identifiable and some not, what now, if anything, ought to be done to rectify these injustices? What obligations do the performers of injustice have toward those whose position is worse than it would have been had the injustice not been done? Or, than it would have been had compensation been paid promptly? How, if at all, do things change if the beneficiaries and those made worse off are not the direct parties in the act of injustice, but, for example, their descendants? Is an injustice done to someone whose holding was itself based upon an unrectified injustice? How far back must one go in wiping clean the historical slate of injustices? What may victims of injustice permissibly do in order to rectify the injustices being done to them, including the many injustices done by persons acting through their government? I do not know of a thorough or theoretically sophisticated treatment of such issues.<sup>2</sup> Idealizing greatly, let us suppose theoretical investigation will produce a principle of rectification. This principle uses historical information about previous situations and injustices done in them (as defined by the first two principles of justice and rights against interference), and information about the actual course of events that flowed from these injustices, until the present, and it yields a description (or descriptions) of holdings in the society. The principle of rectification presumably will make use of its best estimate of subjunctive information about what would have occurred (or a

probability distribution over what might have occurred, using the expected value) if the injustice had not taken place. If the actual description of holdings turns out not to be one of the descriptions yielded by the principle, then one of the descriptions yielded must be realized.\*

The general outlines of the theory of justice in holdings are that the holdings of a person are just if he is entitled to them by the principles of justice in acquisition and transfer, or by the principle of rectification of injustice (as specified by the first two principles). If each person's holdings are just, then the total set (distribution) of holdings is just. To turn these general outlines into a specific theory we would have to specify the details of each of the three principles of justice in holdings: the principle of acquisition of holdings, the principle of transfer of holdings, and the principle of rectification of violations of the first two principles. I shall not attempt that task here. (Locke's principle of justice in acquisition is discussed below.)

#### HISTORICAL PRINCIPLES AND END-RESULT PRINCIPLES

The general outlines of the entitlement theory illuminate the nature and defects of other conceptions of distributive justice. The entitlement theory of justice in distribution is *historical*; whether a distribution is just depends upon how it came about. In contrast, *current time-slice principles* of justice hold that the justice of a distribution is determined by how things are distributed (who has what) as judged by some *structural* principle(s) of just distribution. A utilitarian who judges between any two distributions by seeing

---

\* If the principle of rectification of violations of the first two principles yields more than one description of holdings, then some choice must be made as to which of these is to be realized. Perhaps the sort of considerations about distributive justice and equality that I argue against play a legitimate role in *this* subsidiary choice. Similarly, there may be room for such considerations in deciding which otherwise arbitrary features a statute will embody, when such features are unavoidable because other considerations do not specify a precise line; yet a line must be drawn.

which has the greater sum of utility and, if the sums tie, applies some fixed equality criterion to choose the more equal distribution, would hold a current time-slice principle of justice. As would someone who had a fixed schedule of trade-offs between the sum of happiness and equality. According to a current time-slice principle, all that needs to be looked at, in judging the justice of a distribution, is who ends up with what; in comparing any two distributions one need look only at the matrix presenting the distributions. No further information need be fed into a principle of justice. It is a consequence of such principles of justice that any two structurally identical distributions are equally just. (Two distributions are structurally identical if they present the same profile, but perhaps have different persons occupying the particular slots. My having ten and your having five, and my having five and your having ten are structurally identical distributions.) Welfare economics is the theory of current time-slice principles of justice. The subject is conceived as operating on matrices representing only current information about distribution. This, as well as some of the usual conditions (for example, the choice of distribution is invariant under relabeling of columns), guarantees that welfare economics will be a current time-slice theory, with all of its inadequacies.

Most persons do not accept current time-slice principles as constituting the whole story about distributive shares. They think it relevant in assessing the justice of a situation to consider not only the distribution it embodies, but also how that distribution came about. If some persons are in prison for murder or war crimes, we do not say that to assess the justice of the distribution in the society we must look only at what this person has, and that person has, and that person has, . . . at the current time. We think it relevant to ask whether someone did something so that he *deserved* to be punished, deserved to have a lower share. Most will agree to the relevance of further information with regard to punishments and penalties. Consider also desired things. One traditional socialist view is that workers are entitled to the product and full fruits of their labor; they have earned it; a distribution is unjust if it does not give the workers what they are entitled to. Such entitlements are based upon some past history. No socialist holding this view would find it comforting to be told that because the actual dis-

tribution *A* happens to coincide structurally with the one he desires *D*, *A* therefore is no less just than *D*; it differs only in that the "parasitic" owners of capital receive under *A* what the workers are entitled to under *D*, and the workers receive under *A* what the owners are entitled to under *D*, namely very little. This socialist rightly, in my view, holds onto the notions of earning, producing, entitlement, desert, and so forth, and he rejects current time-slice principles that look only to the structure of the resulting set of holdings. (The set of holdings resulting from what? Isn't it implausible that how holdings are produced and come to exist has no effect at all on who should hold what?) His mistake lies in his view of what entitlements arise out of what sorts of productive processes.

We construe the position we discuss too narrowly by speaking of *current* time-slice principles. Nothing is changed if structural principles operate upon a time sequence of current time-slice profiles and, for example, give someone more now to counterbalance the less he has had earlier. A utilitarian or an egalitarian or any mixture of the two over time will inherit the difficulties of his more myopic comrades. He is not helped by the fact that *some* of the information others consider relevant in assessing a distribution is reflected, unrecoverably, in past matrices. Henceforth, we shall refer to such unhistorical principles of distributive justice, including the current time-slice principles, as *end-result principles* or *end-state principles*.

In contrast to end-result principles of justice, *historical principles* of justice hold that past circumstances or actions of people can create differential entitlements or differential deserts to things. An injustice can be worked by moving from one distribution to another structurally identical one, for the second, in profile the same, may violate people's entitlements or deserts; it may not fit the actual history.

#### PATTERNING

The entitlement principles of justice in holdings that we have sketched are historical principles of justice. To better understand their precise character, we shall distinguish them from another

subclass of the historical principles. Consider, as an example, the principle of distribution according to moral merit. This principle requires that total distributive shares vary directly with moral merit; no person should have a greater share than anyone whose moral merit is greater. (If moral merit could be not merely ordered but measured on an interval or ratio scale, stronger principles could be formulated.) Or consider the principle that results by substituting "usefulness to society" for "moral merit" in the previous principle. Or instead of "distribute according to moral merit," or "distribute according to usefulness to society," we might consider "distribute according to the weighted sum of moral merit, usefulness to society, and need," with the weights of the different dimensions equal. Let us call a principle of distribution *patterned* if it specifies that a distribution is to vary along with some natural dimension, weighted sum of natural dimensions, or lexicographic ordering of natural dimensions. And let us say a distribution is *patterned* if it accords with some patterned principle. (I speak of natural dimensions, admittedly without a general criterion for them, because for any set of holdings some artificial dimensions can be gimmicked up to vary along with the distribution of the set.) The principle of distribution in accordance with moral merit is a patterned historical principle, which specifies a patterned distribution. "Distribute according to I.Q." is a patterned principle that looks to information not contained in distributional matrices. It is not historical, however, in that it does not look to any past actions creating differential entitlements to evaluate a distribution; it requires only distributional matrices whose columns are labeled by I.Q. scores. The distribution in a society, however, may be composed of such simple patterned distributions, without itself being simply patterned. Different sectors may operate different patterns, or some combination of patterns may operate in different proportions across a society. A distribution composed in this manner, from a small number of patterned distributions, we also shall term "patterned." And we extend the use of "pattern" to include the overall designs put forth by combinations of end-state principles.

Almost every suggested principle of distributive justice is patterned: to each according to his moral merit, or needs, or marginal product, or how hard he tries, or the weighted sum of the forego-

ing, and so on. The principle of entitlement we have sketched is *not* patterned.\* There is no one natural dimension or weighted sum or combination of a small number of natural dimensions that yields the distributions generated in accordance with the principle of entitlement. The set of holdings that results when some persons receive their marginal products, others win at gambling, others receive a share of their mate's income, others receive gifts from foundations, others receive interest on loans, others receive gifts from admirers, others receive returns on investment, others make for themselves much of what they have, others find things, and so on, will not be patterned. Heavy strands of patterns will run through it; significant portions of the variance in holdings will be accounted for by pattern-variables. If most people most of the time choose to transfer some of their entitlements to others only in exchange for something from them, then a large part of what many people hold will vary with what they held that others wanted. More details are provided by the theory of marginal productivity. But gifts to relatives, charitable donations, bequests to children, and the like, are not best conceived, in the first instance, in this manner. Ignoring the strands of pattern, let us suppose for the moment that a distribution actually arrived at by the operation of the principle of entitlement is random with respect to any pattern. Though the resulting set of holdings will be unpatterned, it will not be incomprehensible, for it can be seen as arising from the operation of a small number of principles. These principles specify how an initial distribution may arise (the principle of acquisition of holdings) and how distributions may be transformed into others

---

\* One might try to squeeze a patterned conception of distributive justice into the framework of the entitlement conception, by formulating a gimmicky obligatory "principle of transfer" that would lead to the pattern. For example, the principle that if one has more than the mean income one must transfer everything one holds above the mean to persons below the mean so as to bring them up to (but not over) the mean. We can formulate a criterion for a "principle of transfer" to rule out such obligatory transfers, or we can say that no correct principle of transfer, no principle of transfer in a free society will be like this. The former is probably the better course, though the latter also is true.

Alternatively, one might think to make the entitlement conception instantiate a pattern, by using matrix entries that express the relative strength of a person's entitlements as measured by some real-valued function. But even if the limitation to natural dimensions failed to exclude this function, the resulting edifice would *not* capture our system of entitlements to *particular* things.

(the principle of transfer of holdings). The process whereby the set of holdings is generated will be intelligible, though the set of holdings itself that results from this process will be unpatterned.

The writings of F. A. Hayek focus less than is usually done upon what patterning distributive justice requires. Hayek argues that we cannot know enough about each person's situation to distribute to each according to his moral merit (but would justice demand we do so if we did have this knowledge?); and he goes on to say, "our objection is against all attempts to impress upon society a deliberately chosen pattern of distribution, whether it be an order of equality or of inequality."<sup>3</sup> However, Hayek concludes that in a free society there will be distribution in accordance with value rather than moral merit; that is, in accordance with the perceived value of a person's actions and services to others. Despite his rejection of a patterned conception of distributive justice, Hayek himself suggests a pattern he thinks justifiable: distribution in accordance with the perceived benefits given to others, leaving room for the complaint that a free society does not realize exactly this pattern. Stating this patterned strand of a free capitalist society more precisely, we get "To each according to how much he benefits others who have the resources for benefiting those who benefit them." This will seem arbitrary unless some acceptable initial set of holdings is specified, or unless it is held that the operation of the system over time washes out any significant effects from the initial set of holdings. As an example of the latter, if almost anyone would have bought a car from Henry Ford, the supposition that it was an arbitrary matter who held the money then (and so bought) would not place Henry Ford's earnings under a cloud. In any event, *his* coming to hold it is not arbitrary. Distribution according to benefits to others *is* a major patterned strand in a free capitalist society, as Hayek correctly points out, but it is only a strand and does not constitute the whole pattern of a system of entitlements (namely, inheritance, gifts for arbitrary reasons, charity, and so on) or a standard that one should insist a society fit. Will people tolerate for long a system yielding distributions that they believe are unpatterned?<sup>4</sup> No doubt people will not long accept a distribution they believe is *unjust*. People want their society to be and to look just. But must the look of justice reside in a

resulting pattern rather than in the underlying generating principles? We are in no position to conclude that the inhabitants of a society embodying an entitlement conception of justice in holdings will find it unacceptable. Still, it must be granted that were people's reasons for transferring some of their holdings to others always irrational or arbitrary, we would find this disturbing. (Suppose people always determined what holdings they would transfer, and to whom, by using a random device.) We feel more comfortable upholding the justice of an entitlement system if most of the transfers under it are done for reasons. This does not mean necessarily that all deserve what holdings they receive. It means only that there is a purpose or point to someone's transferring a holding to one person rather than to another; that usually we can see what the transferrer thinks he's gaining, what cause he thinks he's serving, what goals he thinks he's helping to achieve, and so forth. Since in a capitalist society people often transfer holdings to others in accordance with how much they perceive these others benefiting them, the fabric constituted by the individual transactions and transfers is largely reasonable and intelligible.\* (Gifts to loved ones, bequests to children, charity to the needy also are nonarbitrary components of the fabric.) In stressing the large strand of distribution in accordance with benefit to others, Hayek shows the point of many transfers, and so shows that the system of transfer of entitlements is not just spinning its gears aimlessly. The system of entitlements is defensible when constituted by the individual aims of individual transactions. No overarching aim is needed, no distributive pattern is required.

To think that the task of a theory of distributive justice is to fill in the blank in "to each according to his \_\_\_\_\_" is to be predis-

---

\* We certainly benefit because great economic incentives operate to get others to spend much time and energy to figure out how to serve us by providing things we will want to pay for. It is not mere paradox mongering to wonder whether capitalism should be criticized for most rewarding and hence encouraging, not individualists like Thoreau who go about their own lives, but people who are occupied with serving others and winning them as customers. But to defend capitalism one need not think businessmen are the finest human types. (I do not mean to join here the general maligning of businessmen, either.) Those who think the finest should acquire the most can try to convince their fellows to transfer resources in accordance with *that* principle.

posed to search for a pattern; and the separate treatment of "from each according to his \_\_\_\_\_" treats production and distribution as two separate and independent issues. On an entitlement view these are *not* two separate questions. Whoever makes something, having bought or contracted for all other held resources used in the process (transferring some of his holdings for these cooperating factors), is entitled to it. The situation is *not* one of something's getting made, and there being an open question of who is to get it. Things come into the world already attached to people having entitlements over them. From the point of view of the historical entitlement conception of justice in holdings, those who start afresh to complete "to each according to his \_\_\_\_\_" treat objects as if they appeared from nowhere, out of nothing. A complete theory of justice might cover this limit case as well; perhaps here is a use for the usual conceptions of distributive justice.<sup>5</sup>

So entrenched are maxims of the usual form that perhaps we should present the entitlement conception as a competitor. Ignoring acquisition and rectification, we might say:

From each according to what he chooses to do, to each according to what he makes for himself (perhaps with the contracted aid of others) and what others choose to do for him and choose to give him of what they've been given previously (under this maxim) and haven't yet expended or transferred.

This, the discerning reader will have noticed, has its defects as a slogan. So as a summary and great simplification (and not as a maxim with any independent meaning) we have:

*From each as they choose, to each as they are chosen.*

#### HOW LIBERTY UPSETS PATTERNS

It is not clear how those holding alternative conceptions of distributive justice can reject the entitlement conception of justice in holdings. For suppose a distribution favored by one of these non-entitlement conceptions is realized. Let us suppose it is your favorite one and let us call this distribution  $D_1$ ; perhaps everyone has an equal share, perhaps shares vary in accordance with some dimen-

sion you treasure. Now suppose that Wilt Chamberlain is greatly in demand by basketball teams, being a great gate attraction. (Also suppose contracts run only for a year, with players being free agents.) He signs the following sort of contract with a team: In each home game, twenty-five cents from the price of each ticket of admission goes to him. (We ignore the question of whether he is "gouging" the owners, letting them look out for themselves.) The season starts, and people cheerfully attend his team's games; they buy their tickets, each time dropping a separate twenty-five cents of their admission price into a special box with Chamberlain's name on it. They are excited about seeing him play; it is worth the total admission price to them. Let us suppose that in one season one million persons attend his home games, and Wilt Chamberlain winds up with \$250,000, a much larger sum than the average income and larger even than anyone else has. Is he entitled to this income? Is this new distribution  $D_2$ , unjust? If so, why? There is *no* question about whether each of the people was entitled to the control over the resources they held in  $D_1$ ; because that was the distribution (your favorite) that (for the purposes of argument) we assumed was acceptable. Each of these persons *chose* to give twenty-five cents of their money to Chamberlain. They could have spent it on going to the movies, or on candy bars, or on copies of *Dissent* magazine, or of *Monthly Review*. But they all, at least one million of them, converged on giving it to Wilt Chamberlain in exchange for watching him play basketball. If  $D_1$  was a just distribution, and people voluntarily moved from it to  $D_2$ , transferring parts of their shares they were given under  $D_1$  (what was it for if not to do something with?), isn't  $D_2$  also just? If the people were entitled to dispose of the resources to which they were entitled (under  $D_1$ ), didn't this include their being entitled to give it to, or exchange it with, Wilt Chamberlain? Can anyone else complain on grounds of justice? Each other person already has his legitimate share under  $D_1$ . Under  $D_1$ , there is nothing that anyone has that anyone else has a claim of justice against. After someone transfers something to Wilt Chamberlain, third parties *still* have their legitimate shares; *their* shares are not changed. By what process could such a transfer among two persons give rise to a legitimate claim of distributive justice on a portion of what was

transferred, by a third party who had no claim of justice on any holding of the others *before* the transfer?\* To cut off objections irrelevant here, we might imagine the exchanges occurring in a socialist society, after hours. After playing whatever basketball he does in his daily work, or doing whatever other daily work he does, Wilt Chamberlain decides to put in *overtime* to earn additional money. (First his work quota is set; he works time over that.) Or imagine it is a skilled juggler people like to see, who puts on shows after hours.

Why might someone work overtime in a society in which it is assumed their needs are satisfied? Perhaps because they care about things other than needs. I like to write in books that I read, and to have easy access to books for browsing at odd hours. It would be very pleasant and convenient to have the resources of Widener Library in my back yard. No society, I assume, will provide such resources close to each person who would like them as part of his regular allotment (under  $D_1$ ). Thus, persons either must do without some extra things that they want, or be allowed to do something extra to get some of these things. On what basis could the inequalities that would eventuate be forbidden? Notice also that small factories would spring up in a socialist society, unless forbidden. I melt down some of my personal possessions (under  $D_1$ ) and build a machine out of the material. I offer you, and others, a philosophy lecture once a week in exchange for your cranking the

---

\* Might not a transfer have instrumental effects on a third party, changing his feasible options? (But what if the two parties to the transfer independently had used their holdings in this fashion?) I discuss this question below, but note here that this question concedes the point for distributions of ultimate intrinsic noninstrumental goods (pure utility experiences, so to speak) that are transferable. It also might be objected that the transfer might make a third party more envious because it worsens his position relative to someone else. I find it incomprehensible how this can be thought to involve a claim of justice. On envy, see Chapter 8.

Here and elsewhere in this chapter, a theory which incorporates elements of pure procedural justice might find what I say acceptable, *if* kept in its proper place; that is, if background institutions exist to ensure the satisfaction of certain conditions on distributive shares. But if these institutions are not themselves the sum or invisible-hand result of people's voluntary (nonaggressive) actions, the constraints they impose require justification. At no point does *our* argument assume any background institutions more extensive than those of the minimal night-watchman state, a state limited to protecting persons against murder, assault, theft, fraud, and so forth.

handle on my machine, whose products I exchange for yet other things, and so on. (The raw materials used by the machine are given to me by others who possess them under  $D_1$ , in exchange for hearing lectures.) Each person might participate to gain things over and above their allotment under  $D_1$ . Some persons even might want to leave their job in socialist industry and work full time in this private sector. I shall say something more about these issues in the next chapter. Here I wish merely to note how private property even in means of production would occur in a socialist society that did not forbid people to use as they wished some of the resources they are given under the socialist distribution  $D_1$ .<sup>6</sup> The socialist society would have to forbid capitalist acts between consenting adults.

The general point illustrated by the Wilt Chamberlain example and the example of the entrepreneur in a socialist society is that no end-state principle or distributional patterned principle of justice can be continuously realized without continuous interference with people's lives. Any favored pattern would be transformed into one unfavored by the principle, by people choosing to act in various ways; for example, by people exchanging goods and services with other people, or giving things to other people, things the transferrers are entitled to under the favored distributional pattern. To maintain a pattern one must either continually interfere to stop people from transferring resources as they wish to, or continually (or periodically) interfere to take from some persons resources that others for some reason chose to transfer to them. (But if some time limit is to be set on how long people may keep resources others voluntarily transfer to them, why let them keep these resources for *any* period of time? Why not have immediate confiscation?) It might be objected that all persons voluntarily will choose to refrain from actions which would upset the pattern. This presupposes unrealistically (1) that all will most want to maintain the pattern (are those who don't, to be "reeducated" or forced to undergo "self-criticism"?), (2) that each can gather enough information about his own actions and the ongoing activities of others to discover which of his actions will upset the pattern, and (3) that diverse and far-flung persons can coordinate their actions to dovetail into the pattern. Compare the manner in which the market is neutral among persons' desires, as it reflects and transmits



widely scattered information via prices, and coordinates persons' activities.

It puts things perhaps a bit too strongly to say that every patterned (or end-state) principle is liable to be thwarted by the voluntary actions of the individual parties transferring some of their shares they receive under the principle. For perhaps some *very* weak patterns are not so thwarted.\* Any distributional pattern with any egalitarian component is overturnable by the voluntary actions of individual persons over time; as is every patterned condition with sufficient content so as actually to have been proposed as presenting the central core of distributive justice. Still, given the possibility that some weak conditions or patterns may not be unstable in this way, it would be better to formulate an explicit description of the kind of interesting and contentful patterns under discussion, and to prove a theorem about their instability. Since the weaker the patterning, the more likely it is that the entitlement system itself satisfies it, a plausible conjecture is that any patterning either is unstable or is satisfied by the entitlement system.

#### SEN'S ARGUMENT

Our conclusions are reinforced by considering a recent general argument of Amartya K. Sen.<sup>7</sup> Suppose individual rights are interpreted as the right to choose which of two alternatives is to be

---

\* Is the patterned principle stable that requires merely that a distribution be Pareto-optimal? One person might give another a gift or bequest that the second could exchange with a third to their mutual benefit. Before the second makes this exchange, there is not Pareto-optimality. Is a stable pattern presented by a principle choosing that among the Pareto-optimal positions that satisfies some further condition *C*? It may seem that there cannot be a counterexample, for won't any voluntary exchange made away from a situation show that the first situation wasn't Pareto-optimal? (Ignore the implausibility of this last claim for the case of bequests.) But principles are to be satisfied over time, during which new possibilities arise. A distribution that at one time satisfies the criterion of Pareto-optimality might not do so when some new possibilities arise (Wilt Chamberlain grows up and starts playing basketball); and though people's activities will tend to move then to a new Pareto-optimal position, *this*

more highly ranked in a social ordering of the alternatives. Add the weak condition that if one alternative unanimously is preferred to another then it is ranked higher by the social ordering. If there are two different individuals each with individual rights, interpreted as above, over different pairs of alternatives (having no members in common), then for some possible preference rankings of the alternatives by the individuals, there is no linear social ordering. For suppose that person *A* has the right to decide among (*X*, *Y*) and person *B* has the right to decide among (*Z*, *W*); and suppose their individual preferences are as follows (and that there are no other individuals). Person *A* prefers *W* to *X* to *Y* to *Z*, and person *B* prefers *Y* to *Z* to *W* to *X*. By the unanimity condition, in the social ordering *W* is preferred to *X* (since each individual prefers it to *X*), and *Y* is preferred to *Z* (since each individual prefers it to *Z*). Also in the social ordering, *X* is preferred to *Y*, by person *A*'s right of choice among these two alternatives. Combining these three binary rankings, we get *W* preferred to *X* preferred to *Y* preferred to *Z*, in the social ordering. However, by person *B*'s right of choice, *Z* must be preferred to *W* in the social ordering. There is no transitive social ordering satisfying all these conditions, and the social ordering, therefore, is nonlinear. Thus far, Sen.

The trouble stems from treating an individual's right to choose among alternatives as the right to determine the relative ordering of these alternatives within a social ordering. The alternative which has individuals rank *pairs* of alternatives, and separately rank the individual alternatives is no better; their ranking of pairs feeds into some method of amalgamating preferences to yield a social ordering of pairs; and the choice among the alternatives in the highest ranked pair in the social ordering is made by the individual with the right to decide between this pair. This system also has the result that an alternative may be selected although *everyone* prefers some other alternative; for example, *A* selects *X* over *Y*, where (*X*, *Y*) somehow is the highest ranked *pair* in the social or-

---

new one need not satisfy the contentful condition *C*. Continual interference will be needed to insure the continual satisfaction of *C*. (The theoretical possibility of a pattern's being maintained by some invisible-hand process that brings it back to an equilibrium that fits the pattern when deviations occur should be investigated.)

dering of pairs, although everyone, including A, prefers *W* to *X*. (But the choice person A was given, however, was only between *X* and *Y*.)

A more appropriate view of individual rights is as follows. Individual rights are co-possible; each person may exercise his rights as he chooses. The exercise of these rights fixes some features of the world. Within the constraints of these fixed features, a choice may be made by a social choice mechanism based upon a social ordering; if there are any choices left to make! Rights do not determine a social ordering but instead set the constraints within which a social choice is to be made, by excluding certain alternatives, fixing others, and so on. (If I have a right to choose to live in New York or in Massachusetts, and I choose Massachusetts, then alternatives involving my living in New York are not appropriate objects to be entered in a social ordering.) Even if all possible alternatives are ordered first, apart from anyone's rights, the situation is not changed: for then the highest ranked alternative *that is not excluded by anyone's exercise of his rights* is instituted. Rights do not determine the position of an alternative or the relative position of two alternatives in a social ordering; they *operate upon* a social ordering to constrain the choice it can yield.

If entitlements to holdings are rights to dispose of them, then social choice must take place *within* the constraints of how people choose to exercise these rights. If any patterning is legitimate, it falls within the domain of social choice, and hence is constrained by people's rights. *How else can one cope with Sen's result?* The alternative of first having a social ranking with rights exercised within its constraints is no alternative at all. Why not just select the top-ranked alternative and forget about rights? If that top-ranked alternative itself leaves some room for individual choice (and here is where "rights" of choice is supposed to enter in) there must be something to stop these choices from transforming it into another alternative. Thus Sen's argument leads us again to the result that patterning requires continuous interference with individuals' actions and choices.<sup>8</sup>

## REDISTRIBUTION AND PROPERTY RIGHTS

Apparently, patterned principles allow people to choose to expend upon themselves, but not upon others, those resources they are entitled to (or rather, receive) under some favored distributional pattern  $D_1$ . For if each of several persons chooses to expend some of his  $D_1$  resources upon one other person, then that other person will receive more than his  $D_1$  share, disturbing the favored distributional pattern. Maintaining a distributional pattern is individualism with a vengeance! Patterned distributional principles do not give people what entitlement principles do, only better distributed. For they do not give the right to choose what to do with what one has; they do not give the right to choose to pursue an end involving (intrinsically, or as a means) the enhancement of another's position. To such views, families are disturbing; for within a family occur transfers that upset the favored distributional pattern. Either families themselves become units to which distribution takes place, the column occupiers (on what rationale?), or loving behavior is forbidden. We should note in passing the ambivalent position of radicals toward the family. Its loving relationships are seen as a model to be emulated and extended across the whole society, at the same time that it is denounced as a suffocating institution to be broken and condemned as a focus of parochial concerns that interfere with achieving radical goals. Need we say that it is not appropriate to enforce across the wider society the relationships of love and care appropriate within a family, relationships which are voluntarily undertaken?\* Incidentally,

---

\* One indication of the stringency of Rawls' difference principle, which we attend to in the second part of this chapter, is its inappropriateness as a governing principle even within a family of individuals who love one another. Should a family devote its resources to maximizing the position of its least well off and least talented child, holding back the other children or using resources for their education and development only if they will follow a policy through their lifetimes of maximizing the position of their least fortunate sibling? Surely not. How then can this even be considered as the appropriate policy for enforcement in the wider society? (I discuss below what I think would be Rawls' reply: that some principles apply at the macro level which do not apply to micro-situations.)

love is an interesting instance of another relationship that is historical, in that (like justice) it depends upon what actually occurred. An adult may come to love another because of the other's characteristics; but it is the other person, and not the characteristics, that is loved.<sup>9</sup> The love is not transferrable to someone else with the same characteristics, even to one who "scores" higher for these characteristics. And the love endures through changes of the characteristics that gave rise to it. One loves the particular person one actually encountered. Why love is historical, attaching to persons in this way and not to characteristics, is an interesting and puzzling question.

Proponents of patterned principles of distributive justice focus upon criteria for determining who is to receive holdings; they consider the reasons for which someone should have something, and also the total picture of holdings. Whether or not it is better to give than to receive, proponents of patterned principles ignore giving altogether. In considering the distribution of goods, income, and so forth, their theories are theories of recipient justice; they completely ignore any right a person might have to give something to someone. Even in exchanges where each party is simultaneously giver and recipient, patterned principles of justice focus only upon the recipient role and its supposed rights. Thus discussions tend to focus on whether people (should) have a right to inherit, rather than on whether people (should) have a right to bequeath or on whether persons who have a right to hold also have a right to choose that others hold in their place. I lack a good explanation of why the usual theories of distributive justice are so recipient oriented; ignoring givers and transferrers and their rights is of a piece with ignoring producers and their entitlements. But why is it *all* ignored?

Patterned principles of distributive justice necessitate *redistributive* activities. The likelihood is small that any actual freely-arrived-at set of holdings fits a given pattern; and the likelihood is nil that it will continue to fit the pattern as people exchange and give. From the point of view of an entitlement theory, redistribution is a serious matter indeed, involving, as it does, the violation of people's rights. (An exception is those takings that fall under the principle of the rectification of injustices.) From other points of view, also, it is serious.

Taxation of earnings from labor is on a par with forced labor.\* Some persons find this claim obviously true: taking the earnings of  $n$  hours labor is like taking  $n$  hours from the person; it is like forcing the person to work  $n$  hours for another's purpose. Others find the claim absurd. But even these, *if* they object to forced labor, would oppose forcing unemployed hippies to work for the benefit of the needy.<sup>†</sup> And they would also object to forcing each person to work five extra hours each week for the benefit of the needy. But a system that takes five hours' wages in taxes does not seem to them like one that forces someone to work five hours, since it offers the person forced a wider range of choice in activities than does taxation in kind with the particular labor specified. (But we can imagine a gradation of systems of forced labor, from one that specifies a particular activity, to one that gives a choice among two activities, to . . . ; and so on up.) Furthermore, people envisage a system with something like a proportional tax on everything above the amount necessary for basic needs. Some think this does not force someone to work extra hours, since there is no fixed number of extra hours he is forced to work, and since he can avoid the tax entirely by earning only enough to cover his basic needs. This is a very uncharacteristic view of forcing for those who *also* think people are forced to do something *whenever* the alternatives they face are considerably worse. However, *neither* view is correct. The fact that others intentionally intervene, in violation of a side constraint against aggression, to threaten force to limit the alternatives, in this case to paying taxes or (presumably the worse alternative) bare subsistence, makes the taxation system one of forced labor and distinguishes it from other cases of limited choices which are not forcings.<sup>10</sup>

\* I am unsure as to whether the arguments I present below show that such taxation merely *is* forced labor; so that "is on a par with" means "is one kind of." Or alternatively, whether the arguments emphasize the great similarities between such taxation and forced labor, to show it is plausible and illuminating to view such taxation in the light of forced labor. This latter approach would remind one of how John Wisdom conceives of the claims of metaphysicians.

† Nothing hangs on the fact that here and elsewhere I speak loosely of *needs*, since I go on, each time, to reject the criterion of justice which includes it. If, however, something did depend upon the notion, one would want to examine it more carefully. For a skeptical view, see Kenneth Minogue, *The Liberal Mind*, (New York: Random House, 1963), pp. 103-112.

The man who chooses to work longer to gain an income more than sufficient for his basic needs prefers some extra goods or services to the leisure and activities he could perform during the possible nonworking hours; whereas the man who chooses not to work the extra time prefers the leisure activities to the extra goods or services he could acquire by working more. Given this, if it would be illegitimate for a tax system to seize some of a man's leisure (forced labor) for the purpose of serving the needy, how can it be legitimate for a tax system to seize some of a man's goods for that purpose? Why should we treat the man whose happiness requires certain material goods or services differently from the man whose preferences and desires make such goods unnecessary for his happiness? Why should the man who prefers seeing a movie (and who has to earn money for a ticket) be open to the required call to aid the needy, while the person who prefers looking at a sunset (and hence need earn no extra money) is not? Indeed, isn't it surprising that redistributionists choose to ignore the man whose pleasures are so easily attainable without extra labor, while adding yet another burden to the poor unfortunate who must work for his pleasures? If anything, one would have expected the reverse. Why is the person with the nonmaterial or nonconsumption desire allowed to proceed unimpeded to his most favored feasible alternative, whereas the man whose pleasures or desires involve material things and who must work for extra money (thereby serving whomever considers his activities valuable enough to pay him) is constrained in what he can realize? Perhaps there is no difference in principle. And perhaps some think the answer concerns merely administrative convenience. (These questions and issues will not disturb those who think that forced labor to serve the needy or to realize some favored end-state pattern is acceptable.) In a fuller discussion we would have (and want) to extend our argument to include interest, entrepreneurial profits, and so on. Those who doubt that this extension can be carried through, and who draw the line here at taxation of income from labor, will have to state rather complicated patterned *historical* principles of distributive justice, since end-state principles would not distinguish *sources* of income in any way. It is enough for now to get away from end-state principles and to make clear how various patterned principles are dependent upon particular views about the sources or the ille-

gitimacy or the lesser legitimacy of profits, interest, and so on; which particular views may well be mistaken.

What sort of right over others does a legally institutionalized end-state pattern give one? The central core of the notion of a property right in *X*, relative to which other parts of the notion are to be explained, is the right to determine what shall be done with *X*; the right to choose which of the constrained set of options concerning *X* shall be realized or attempted.<sup>11</sup> The constraints are set by other principles or laws operating in the society; in our theory, by the Lockean rights people possess (under the minimal state). My property rights in my knife allow me to leave it where I will, but not in your chest. I may choose which of the acceptable options involving the knife is to be realized. This notion of property helps us to understand why earlier theorists spoke of people as having property in themselves and their labor. They viewed each person as having a right to decide what would become of himself and what he would do, and as having a right to reap the benefits of what he did.

This right of selecting the alternative to be realized from the constrained set of alternatives may be held by an *individual* or by a *group* with some procedure for reaching a joint decision; or the right may be passed back and forth, so that one year I decide what's to become of *X*, and the next year you do (with the alternative of destruction, perhaps, being excluded). Or, during the same time period, some types of decisions about *X* may be made by me, and others by you. And so on. We lack an adequate, fruitful, analytical apparatus for classifying the *types* of constraints on the set of options among which choices are to be made, and the *types* of ways decision powers can be held, divided, and amalgamated. A *theory* of property would, among other things, contain such a classification of constraints and decision modes, and from a small number of principles would follow a host of interesting statements about the *consequences* and effects of certain combinations of constraints and modes of decision.

When end-result principles of distributive justice are built into the legal structure of a society, they (as do most patterned principles) give each citizen an enforceable claim to some portion of the total social product; that is, to some portion of the sum total of the individually and jointly made products. This total product is

produced by individuals laboring, using means of production others have saved to bring into existence, by people organizing production or creating means to produce new things or things in a new way. It is on this batch of individual activities that patterned distributional principles give each individual an enforceable claim. Each person has a claim to the activities and the products of other persons, independently of whether the other persons enter into particular relationships that give rise to these claims, and independently of whether they voluntarily take these claims upon themselves, in charity or in exchange for something.

Whether it is done through taxation on wages or on wages over a certain amount, or through seizure of profits, or through there being a big *social pot* so that it's not clear what's coming from where and what's going where, patterned principles of distributive justice involve appropriating the actions of other persons. Seizing the results of someone's labor is equivalent to seizing hours from him and directing him to carry on various activities. If people force you to do certain work, or unrewarded work, for a certain period of time, they decide what you are to do and what purposes your work is to serve apart from your decisions. This process whereby they take this decision from you makes them a *part-owner* of you; it gives them a property right in you. Just as having such partial control and power of decision, by right, over an animal or inanimate object would be to have a property right in it.

End-state and most patterned principles of distributive justice institute (partial) ownership by others of people and their actions and labor. These principles involve a shift from the classical liberals' notion of self-ownership to a notion of (partial) property rights in *other* people.

Considerations such as these confront end-state and other patterned conceptions of justice with the question of whether the actions necessary to achieve the selected pattern don't themselves violate moral side constraints. Any view holding that there are moral side constraints on actions, that not all moral considerations can be built into end states that are to be achieved (see Chapter 3, pp. 28–30), must face the possibility that some of its goals are not achievable by any morally permissible available means. An entitlement theorist will face such conflicts in a society that deviates from the principles of justice for the generation of holdings, if and

only if the only actions available to realize the principles themselves violate some moral constraints. Since deviation from the first two principles of justice (in acquisition and transfer) will involve other persons' direct and aggressive intervention to violate rights, and since moral constraints will not exclude defensive or retributive action in such cases, the entitlement theorist's problem rarely will be pressing. And whatever difficulties he has in applying the principle of rectification to persons who did not themselves violate the first two principles are difficulties in balancing the conflicting considerations so as correctly to formulate the complex principle of rectification itself; he will not violate moral side constraints by applying the principle. Proponents of patterned conceptions of justice, however, often will face head-on clashes (and poignant ones if they cherish each party to the clash) between moral side constraints on how individuals may be treated and their patterned conception of justice that presents an end state or other pattern that *must* be realized.

May a person emigrate from a nation that has institutionalized some end-state or patterned distributional principle? For some principles (for example, Hayek's) emigration presents no theoretical problem. But for others it is a tricky matter. Consider a nation having a compulsory scheme of minimal social provision to aid the neediest (or one organized so as to maximize the position of the worst-off group); no one may opt out of participating in it. (None may say, "Don't compel me to contribute to others and don't provide for me via this compulsory mechanism if I am in need.") Everyone above a certain level is forced to contribute to aid the needy. But if emigration from the country were allowed, anyone could choose to move to another country that did not have compulsory social provision but otherwise was (as much as possible) identical. In such a case, the person's *only* motive for leaving would be to avoid participating in the compulsory scheme of social provision. And if he does leave, the needy in his initial country will receive no (compelled) help from him. What rationale yields the result that the person be permitted to emigrate, yet forbidden to stay and opt out of the compulsory scheme of social provision? If providing for the needy is of overriding importance, this does militate against allowing internal opting out; but it also speaks against allowing external emigration. (Would it also support, to

some extent, the kidnapping of persons living in a place without compulsory social provision, who could be forced to make a contribution to the needy in your community?) Perhaps the crucial component of the position that allows emigration solely to avoid certain arrangements, while not allowing anyone internally to opt out of them, is a concern for fraternal feelings within the country. "We don't want anyone here who doesn't contribute, who doesn't care enough about the others to contribute." That concern, in this case, would have to be tied to the view that forced aiding tends to produce fraternal feelings between the aided and the aider (or perhaps merely to the view that the knowledge that someone or other voluntarily is not aiding produces unfraternal feelings).

#### LOCKE'S THEORY OF ACQUISITION

Before we turn to consider other theories of justice in detail, we must introduce an additional bit of complexity into the structure of the entitlement theory. This is best approached by considering Locke's attempt to specify a principle of justice in acquisition. Locke views property rights in an unowned object as originating through someone's mixing his labor with it. This gives rise to many questions. What are the boundaries of what labor is mixed with? If a private astronaut clears a place on Mars, has he mixed his labor with (so that he comes to own) the whole planet, the whole uninhabited universe, or just a particular plot? Which plot does an act bring under ownership? The minimal (possibly disconnected) area such that an act decreases entropy in that area, and not elsewhere? Can virgin land (for the purposes of ecological investigation by high-flying airplane) come under ownership by a Lockean process? Building a fence around a territory presumably would make one the owner of only the fence (and the land immediately underneath it).

Why does mixing one's labor with something make one the owner of it? Perhaps because one owns one's labor, and so one comes to own a previously unowned thing that becomes permeated with what one owns. Ownership seeps over into the rest. But why isn't mixing what I own with what I don't own a way of losing

what I own rather than a way of gaining what I don't? If I own a can of tomato juice and spill it in the sea so that its molecules (made radioactive, so I can check this) mingle evenly throughout the sea, do I thereby come to own the sea, or have I foolishly dissipated my tomato juice? Perhaps the idea, instead, is that laboring on something improves it and makes it more valuable; and anyone is entitled to own a thing whose value he has created. (Reinforcing this, perhaps, is the view that laboring is unpleasant. If some people made things effortlessly, as the cartoon characters in *The Yellow Submarine* trail flowers in their wake, would they have lesser claim to their own products whose making didn't *cost* them anything?) Ignore the fact that laboring on something may make it less valuable (spraying pink enamel paint on a piece of driftwood that you have found). Why should one's entitlement extend to the whole object rather than just to the *added value* one's labor has produced? (Such reference to value might also serve to delimit the extent of ownership; for example, substitute "increases the value of" for "decreases entropy in" in the above entropy criterion.) No workable or coherent value-added property scheme has yet been devised, and any such scheme presumably would fall to objections (similar to those) that fell the theory of Henry George.

It will be implausible to view improving an object as giving full ownership to it, if the stock of unowned objects that might be improved is limited. For an object's coming under one person's ownership changes the situation of all others. Whereas previously they were at liberty (in Hohfeld's sense) to use the object, they now no longer are. This change in the situation of others (by removing their liberty to act on a previously unowned object) need not worsen their situation. If I appropriate a grain of sand from Coney Island, no one else may now do as they will with *that* grain of sand. But there are plenty of other grains of sand left for them to do the same with. Or if not grains of sand, then other things. Alternatively, the things I do with the grain of sand I appropriate might improve the position of others, counterbalancing their loss of the liberty to use that grain. The crucial point is whether appropriation of an unowned object worsens the situation of others.

Locke's proviso that there be "enough and as good left in common for others" (sect. 27) is meant to ensure that the situation of others is not worsened. (If this proviso is met is there any motiva-

tion for his further condition of nonwaste?) It is often said that this proviso once held but now no longer does. But there appears to be an argument for the conclusion that if the proviso no longer holds, then it cannot ever have held so as to yield permanent and inheritable property rights. Consider the first person *Z* for whom there is not enough and as good left to appropriate. The last person *Y* to appropriate left *Z* without his previous liberty to act on an object, and so worsened *Z*'s situation. So *Y*'s appropriation is not allowed under Locke's proviso. Therefore the next to last person *X* to appropriate left *Y* in a worse position, for *X*'s act ended permissible appropriation. Therefore *X*'s appropriation wasn't permissible. But then the appropriator two from last, *W*, ended permissible appropriation and so, since it worsened *X*'s position, *W*'s appropriation wasn't permissible. And so on back to the first person *A* to appropriate a permanent property right.

This argument, however, proceeds too quickly. Someone may be made worse off by another's appropriation in two ways: first, by losing the opportunity to improve his situation by a particular appropriation or any one; and second, by no longer being able to use freely (without appropriation) what he previously could. A *stringent* requirement that another not be made worse off by an appropriation would exclude the first way if nothing else counterbalances the diminution in opportunity, as well as the second. A *weaker* requirement would exclude the second way, though not the first. With the weaker requirement, we cannot zip back so quickly from *Z* to *A*, as in the above argument; for though person *Z* can no longer *appropriate*, there may remain some for him to *use* as before. In this case *Y*'s appropriation would not violate the weaker Lockean condition. (With less remaining that people are at liberty to use, users might face more inconvenience, crowding, and so on; in that way the situation of others might be worsened, unless appropriation stopped far short of such a point.) It is arguable that no one legitimately can complain if the weaker provision is satisfied. However, since this is less clear than in the case of the more stringent proviso, Locke may have intended this stringent proviso by "enough and as good" remaining, and perhaps he meant the non-waste condition to delay the end point from which the argument zips back.

Is the situation of persons who are unable to appropriate (there being no more accessible and useful unowned objects) worsened by a system allowing appropriation and permanent property? Here enter the various familiar social considerations favoring private property: it increases the social product by putting means of production in the hands of those who can use them most efficiently (profitably); experimentation is encouraged, because with separate persons controlling resources, there is no one person or small group whom someone with a new idea must convince to try it out; private property enables people to decide on the pattern and types of risks they wish to bear, leading to specialized types of risk bearing; private property protects future persons by leading some to hold back resources from current consumption for future markets; it provides alternate sources of employment for unpopular persons who don't have to convince any one person or small group to hire them, and so on. These considerations enter a Lockean theory to support the claim that appropriation of private property satisfies the intent behind the "enough and as good left over" proviso, *not* as a utilitarian justification of property. They enter to rebut the claim that because the proviso is violated no natural right to private property can arise by a Lockean process. The difficulty in working such an argument to show that the proviso is satisfied is in fixing the appropriate base line for comparison. Lockean appropriation makes people no worse off than they would be *how?*<sup>12</sup> This question of fixing the baseline needs more detailed investigation than we are able to give it here. It would be desirable to have an estimate of the general economic importance of original appropriation in order to see how much leeway there is for differing theories of appropriation and of the location of the baseline. Perhaps this importance can be measured by the percentage of all income that is based upon untransformed raw materials and given resources (rather than upon human actions), mainly rental income representing the unimproved value of land, and the price of raw material *in situ*, and by the percentage of current wealth which represents such income in the past.\*

\* I have not seen a precise estimate. David Friedman, *The Machinery of Freedom* (N.Y.: Harper & Row, 1973), pp. xiv, xv, discusses this issue and sug-

We should note that it is not only persons favoring *private* property who need a theory of how property rights legitimately originate. Those believing in collective property, for example those believing that a group of persons living in an area jointly own the territory, or its mineral resources, also must provide a theory of how such property rights arise; they must show why the persons living there have rights to determine what is done with the land and resources there that persons living elsewhere don't have (with regard to the same land and resources).

#### THE PROVISIO

Whether or not Locke's particular theory of appropriation can be spelled out so as to handle various difficulties, I assume that any adequate theory of justice in acquisition will contain a proviso similar to the weaker of the ones we have attributed to Locke. A process normally giving rise to a permanent bequeathable property right in a previously unowned thing will not do so if the position of others no longer at liberty to use the thing is thereby worsened. It is important to specify *this* particular mode of worsening the situation of others, for the proviso does not encompass other modes. It does not include the worsening due to more limited opportunities to appropriate (the first way above, corresponding to the more stringent condition), and it does not include how I "worsen" a seller's position if I appropriate materials to make some of what he is selling, and then enter into competition with him. Someone whose appropriation otherwise would violate the proviso still may appropriate provided he compensates the others so that their situation is not thereby worsened; unless he does compensate these others, his appropriation will violate the proviso of the principle of justice in acquisition and will be an illegitimate one.\* A theory of

---

gests 5 percent of U.S. national income as an upper limit for the first two factors mentioned. However he does not attempt to estimate the percentage of current wealth which is based upon such income in the past. (The vague notion of "based upon" merely indicates a topic needing investigation.)

\* Fourier held that since the process of civilization had deprived the members of society of certain liberties (to gather, pasture, engage in the chase),

appropriation incorporating this Lockean proviso will handle correctly the cases (objections to the theory lacking the proviso) where someone appropriates the total supply of something necessary for life.\*

A theory which includes this proviso in its principle of justice in acquisition must also contain a more complex principle of justice in transfer. Some reflection of the proviso about appropriation constrains later actions. If my appropriating all of a certain substance violates the Lockean proviso, then so does my appropriating some and purchasing all the rest from others who obtained it without otherwise violating the Lockean proviso. If the proviso excludes someone's appropriating all the drinkable water in the world, it also excludes his purchasing it all. (More weakly, and messily, it may exclude his charging certain prices for some of his supply.) This proviso (almost?) never will come into effect; the more someone acquires of a scarce substance which others want, the higher the price of the rest will go, and the more difficult it will become for him to acquire it all. But still, we can imagine, at least, that something like this occurs: someone makes simulta-

---

a socially guaranteed minimum provision for persons was justified as compensation for the loss (Alexander Gray, *The Socialist Tradition* (New York: Harper & Row, 1968), p. 188). But this puts the point too strongly. This compensation would be due those persons, if any, for whom the process of civilization was a *net loss*, for whom the benefits of civilization did not counterbalance being deprived of these particular liberties.

\* For example, Rashdall's case of someone who comes upon the only water in the desert several miles ahead of others who also will come to it and appropriates it all. Hastings Rashdall, "The Philosophical Theory of Property," in *Property, its Duties and Rights* (London: MacMillan, 1915).

We should note Ayn Rand's theory of property rights ("Man's Rights" in *The Virtue of Selfishness* (New York: New American Library, 1964), p. 94), wherein these follow from the right to life, since people need physical things to live. But a right to life is not a right to whatever one needs to live; other people may have rights over these other things (see Chapter 3 of this book). At most, a right to life would be a right to have or strive for whatever one needs to live, provided that having it does not violate anyone else's rights. With regard to material things, the question is whether having it does violate any right of others. (Would appropriation of all unowned things do so? Would appropriating the water hole in Rashdall's example?) Since special considerations (such as the Lockean proviso) may enter with regard to material property, one *first* needs a theory of property rights before one can apply any supposed right to life (as amended above). Therefore the right to life cannot provide the foundation for a theory of property rights.



neous secret bids to the separate owners of a substance, each of whom sells assuming he can easily purchase more from the other owners; or some natural catastrophe destroys all of the supply of something except that in one person's possession. The total supply could not be permissibly appropriated by one person at the beginning. His later acquisition of it all does not show that the original appropriation violated the proviso (even by a reverse argument similar to the one above that tried to zip back from Z to A). Rather, it is the combination of the original appropriation *plus* all the later transfers and actions that violates the Lockean proviso.

Each owner's title to his holding includes the historical shadow of the Lockean proviso on appropriation. This excludes his transferring it into an agglomeration that does violate the Lockean proviso and excludes his using it in a way, in coordination with others or independently of them, so as to violate the proviso by making the situation of others worse than their baseline situation. Once it is known that someone's ownership runs afoul of the Lockean proviso, there are stringent limits on what he may do with (what it is difficult any longer unreservedly to call) "his property." Thus a person may not appropriate the only water hole in a desert and charge what he will. Nor may he charge what he will if he possesses one, and unfortunately it happens that all the water holes in the desert dry up, except for his. This unfortunate circumstance, admittedly no fault of his, brings into operation the Lockean proviso and limits his property rights.\* Similarly, an owner's property right in the only island in an area does not allow him to order a castaway from a shipwreck off his island as a trespasser, for this would violate the Lockean proviso.

Notice that the theory does not say that owners do have these rights, but that the rights are overridden to avoid some catastrophe. (Overridden rights do not disappear; they leave a trace of a sort absent in the cases under discussion.)<sup>13</sup> There is no such external (and *ad hoc*?) overriding. Considerations internal to the theory of property itself, to its theory of acquisition and appropria-

\* The situation would be different if his water hole didn't dry up, due to special precautions he took to prevent this. Compare our discussion of the case in the text with Hayek, *The Constitution of Liberty*, p. 136; and also with Ronald Hamowy, "Hayek's Concept of Freedom; A Critique," *New Individualist Review*, April 1961, pp. 28-31.

tion, provide the means for handling such cases. The results, however, may be coextensive with some condition about catastrophe, since the baseline for comparison is so low as compared to the productiveness of a society with private appropriation that the question of the Lockean proviso being violated arises only in the case of catastrophe (or a desert-island situation).

The fact that someone owns the total supply of something necessary for others to stay alive does *not* entail that his (or anyone's) appropriation of anything left some people (immediately or later) in a situation worse than the baseline one. A medical researcher who synthesizes a new substance that effectively treats a certain disease and who refuses to sell except on his terms does not worsen the situation of others by depriving them of whatever he has appropriated. The others easily can possess the same materials he appropriated; the researcher's appropriation or purchase of chemicals didn't make those chemicals scarce in a way so as to violate the Lockean proviso. Nor would someone else's purchasing the total supply of the synthesized substance from the medical researcher. The fact that the medical researcher uses easily available chemicals to synthesize the drug no more violates the Lockean proviso than does the fact that the only surgeon able to perform a particular operation eats easily obtainable food in order to stay alive and to have the energy to work. This shows that the Lockean proviso is not an "end-state principle"; it focuses on a particular way that appropriative actions affect others, and not on the structure of the situation that results.<sup>14</sup>

Intermediate between someone who takes all of the public supply and someone who makes the total supply out of easily obtainable substances is someone who appropriates the total supply of something in a way that does not deprive the others of it. For example, someone finds a new substance in an out-of-the-way place. He discovers that it effectively treats a certain disease and appropriates the total supply. He does not worsen the situation of others; if he did not stumble upon the substance no one else would have, and the others would remain without it. However, as time passes, the likelihood increases that others would have come across the substance; upon this fact might be based a limit to his property right in the substance so that others are not below their baseline position; for example, its bequest might be limited. The

theme of someone worsening another's situation by depriving him of something he otherwise would possess may also illuminate the example of patents. An inventor's patent does not deprive others of an object which would not exist if not for the inventor. Yet patents would have this effect on others who independently invent the object. Therefore, these independent inventors, upon whom the burden of proving independent discovery may rest, should not be excluded from utilizing their own invention as they wish (including selling it to others). Furthermore, a known inventor drastically lessens the chances of actual independent invention. For persons who know of an invention usually will not try to reinvent it, and the notion of independent discovery here would be murky at best. Yet we may assume that in the absence of the original invention, sometime later someone else would have come up with it. This suggests placing a time limit on patents, as a rough rule of thumb to approximate how long it would have taken, in the absence of knowledge of the invention, for independent discovery.

I believe that the free operation of a market system will not actually run afoul of the Lockean proviso. (Recall that crucial to our story in Part I of how a protective agency becomes dominant and a *de facto* monopoly is the fact that it wields force in situations of conflict, and is not merely in competition, with other agencies. A similar tale cannot be told about other businesses.) If this is correct, the proviso will not play a very important role in the activities of protective agencies and will not provide a significant opportunity for future state action. Indeed, were it not for the effects of previous *illegitimate* state action, people would not think the possibility of the proviso's being violated as of more interest than any other logical possibility. (Here I make an empirical historical claim; as does someone who disagrees with this.) This completes our indication of the complication in the entitlement theory introduced by the Lockean proviso.

## SECTION II

## RAWLS' THEORY

We can bring our discussion of distributive justice into sharper focus by considering in some detail John Rawls' recent contribution to the subject. *A Theory of Justice*<sup>15</sup> is a powerful, deep, subtle, wide-ranging, systematic work in political and moral philosophy which has not seen its like since the writings of John Stuart Mill, if then. It is a fountain of illuminating ideas, integrated together into a lovely whole. Political philosophers now must either work within Rawls' theory or explain why not. The considerations and distinctions we have developed are illuminated by, and help illuminate, Rawls' masterful presentation of an alternative conception. Even those who remain unconvinced after wrestling with Rawls' systematic vision will learn much from closely studying it. I do not speak only of the Millian sharpening of one's views in combating (what one takes to be) error. It is impossible to read Rawls' book without incorporating much, perhaps transmuted, into one's own deepened view. And it is impossible to finish his book without a new and inspiring vision of what a moral theory may attempt to do and unite; of how *beautiful* a whole theory can be. I permit myself to concentrate here on disagreements with Rawls only because I am confident that my readers will have discovered for themselves its many virtues.

## SOCIAL COOPERATION

I shall begin by considering the role of the principles of justice. Let us assume, to fix ideas, that a society is a more or less self-sufficient association of persons who in their relations to one another recognize certain rules of conduct as binding and who for the most part act in accordance with them. Suppose further that these rules specify a system of cooperation designed to advance the good of those taking part in it. Then, although a society is a cooperative venture for mutual advantage, it is

typically marked by a conflict as well as by an identity of interests. There is an identity of interests since social cooperation makes possible a better life for all than any would have if each were to live solely by his own efforts. There is a conflict of interests since persons are not indifferent as to how the greater benefits produced by their collaboration are distributed, for in order to pursue their ends they each prefer a larger to a lesser share. A set of principles is required for choosing among the various social arrangements which determine this division of advantages and for underwriting an agreement on the proper distributive shares. These principles are the principles of social justice: they provide a way of assigning rights and duties in the basic institutions of society and they define the appropriate distribution of the benefits and burdens of social cooperation.<sup>16</sup>

Let us imagine  $n$  individuals who do not cooperate together and who each live solely by their own efforts. Each person  $i$  receives a payoff, return, income, and so forth,  $S_i$ ; the sum total of what each individual gets acting separately is

$$S = \sum_{i=1}^n S_i.$$

By cooperating together they can obtain a larger sum total  $T$ . The problem of distributive social justice, according to Rawls, is how these benefits of cooperation are to be distributed or allocated. This problem might be conceived of in two ways: how is the total  $T$  to be allocated? Or, how is the incremental amount due to social cooperation, that is the benefits of social cooperation  $T - S$ , to be allocated? The latter formulation assumes that each individual  $i$  receives from the subtotal  $S$  of  $T$ , his share  $S_i$ . The two statements of the problem differ. When combined with the noncooperative distribution of  $S$  (each  $i$  getting  $S_i$ ), a "fair-looking" distribution of  $T - S$  under the second version may not yield a "fair-looking" distribution of  $T$  (the first version). Alternatively, a fair-looking distribution of  $T$  may give a particular individual  $i$  less than his share  $S_i$ . (The constraint  $T_i \geq S_i$  on the answer to the first formulation of the problem, where  $T_i$  is the share in  $T$  of the  $i^{\text{th}}$  individual, would exclude this possibility.) Rawls, without distinguishing these two formulations of the problem, writes as though his concern is the first one, that is, how the total sum  $T$  is to be distributed. One might claim, to support a focus on the first issue, that due to the enormous benefits of social cooperation, the noncooperative shares  $S_i$  are so small in comparison to any cooperative ones  $T_i$  that they

may be ignored in setting up the problem of social justice. Though we should note that this certainly is not how people entering into cooperation with one another would agree to conceive of the problem of dividing up cooperation's benefits.

Why does social cooperation *create* the problem of distributive justice? Would there be no problem of justice and no need for a theory of justice, if there was no social cooperation at all, if each person got his share solely by his own efforts? If we suppose, as Rawls seems to, that this situation does *not* raise questions of distributive justice, then in virtue of what facts about social cooperation do these questions of justice emerge? What is it about social cooperation that gives rise to issues of justice? It cannot be said that there will be conflicting claims only where there is social cooperation; that individuals who produce independently and (initially) fend for themselves will not make claims of justice on each other. If there were ten Robinson Crusoes, each working alone for two years on separate islands, who discovered each other and the facts of their different allotments by radio communication via transmitters left twenty years earlier, could they not make claims on each other, supposing it were possible to transfer goods from one island to the next? <sup>17</sup> Wouldn't the one with least make a claim on ground of need, or on the ground that his island was naturally poorest, or on the ground that he was naturally least capable of fending for himself? Mightn't he say that justice demanded he be given some more by the others, claiming it unfair that he should receive so much less and perhaps be destitute, perhaps starving? He might go on to say that the different individual noncooperative shares stem from differential natural endowments, which are not deserved, and that the task of justice is to rectify these arbitrary facts and inequities. Rather than its being the case that no one *will* make such claims in the situation lacking social cooperation, perhaps the point is that such claims clearly would be without merit. Why would they clearly be without merit? In the social noncooperation situation, it might be said, each individual deserves what he gets unaided by his own efforts; or rather, no one else can make a claim *of justice* against this holding. It is pellucidly clear in this situation who is entitled to what, so no theory of justice is needed. On this view social cooperation introduces a muddying of the waters that makes it unclear or indeterminate who is

entitled to what. Rather than saying that no theory of justice applies to this noncooperative case, (wouldn't it be unjust if someone stole another's products in the noncooperative situation?), I would say that it is a clear case of application of the correct theory of justice: the entitlement theory.

How does social cooperation change things so that the same entitlement principles that apply to the noncooperative cases become inapplicable or inappropriate to cooperative ones? It might be said that one cannot disentangle the contributions of distinct individuals who cooperate; everything is everyone's joint product. On this joint product, or on any portion of it, each person plausibly will make claims of equal strength; all have an equally good claim, or at any rate no person has a distinctly better claim than any other. Somehow (this line of thought continues), it must be decided how this total product of joint social cooperation (to which individual entitlements do not apply differentially) is to be divided up: this is the problem of distributive justice.

Don't individual entitlements apply to parts of the cooperatively produced product? First, suppose that social cooperation is based upon division of labor, specialization, comparative advantage, and exchange; each person works singly to transform some input he receives, contracting with others who further transform or transport his product until it reaches its ultimate consumer. People cooperate in making things but they work separately; each person is a miniature firm.<sup>18</sup> The products of each person are easily identifiable, and exchanges are made in open markets with prices set competitively, given informational constraints, and so forth. In such a system of social cooperation, what is the task of a theory of justice? It might be said that whatever holdings result will depend upon the exchange ratios or prices at which exchanges are made, and therefore that the task of a theory of justice is to set criteria for "fair prices." This is hardly the place to trace the serpentine windings of theories of a just price. It is difficult to see why these issues should even arise here. People are choosing to make exchanges with other people and to transfer entitlements, with no restrictions on their freedom to trade with any other party at any mutually acceptable ratio.<sup>19</sup> Why does such sequential social cooperation, linked together by people's voluntary exchanges, raise any special problems about how things are to be distributed? Why isn't the

appropriate (a not inappropriate) set of holdings just the one which *actually occurs* via this process of mutually-agreed-to exchanges whereby people choose to give to others what they are entitled to give or hold?

Let us now drop our assumption that people work independently, cooperating only in sequence via voluntary exchanges, and instead consider people who work together jointly to produce something. Is it now impossible to disentangle people's respective contributions? The question here is not whether marginal productivity theory is an appropriate theory of fair or just shares, but whether there is some coherent notion of identifiable marginal product. It seems unlikely that Rawls' theory rests on the strong claim that there is no such reasonably serviceable notion. Anyway, once again we have a situation of a large number of bilateral exchanges: owners of resources reaching separate agreements with entrepreneurs about the use of their resources, entrepreneurs reaching agreements with individual workers, or groups of workers first reaching some joint agreement and then presenting a package to an entrepreneur, and so forth. People transfer their holdings or labor in free markets, with the exchange ratios (prices) determined in the usual manner. If marginal productivity theory is reasonably adequate, people will be receiving, in these voluntary transfers of holdings, roughly their marginal products.\*

---

\* Receiving this, we should note, is not the same as receiving the equivalent of what the person *causes* to exist, or *produces*. The marginal product of a unit of  $F_1$  with respect to factor  $F_2, \dots, F_n$  is a *subjunctive* notion; it is the difference between the total product of  $F_1, \dots, F_n$  used most efficiently (as efficiently as known how, given prudence about many costs in finding out the most efficient use of factors) and the total product of the most efficient use of  $F_2, \dots, F_n$  along with a unit less of  $F_1$ . But these two different most efficient uses of  $F_2, \dots, F_n$  along with a unit less of  $F_1$  (one with the additional unit of  $F_1$ , the other without it) will use them differently. And  $F_1$ 's marginal product (with respect to the other factors), what everyone reasonably would pay for an additional unit of  $F_1$ , will not be what it *causes* (it causes) combined with  $F_2, \dots, F_n$  and the other units of  $F_1$ , but rather the difference it makes, the difference there would be if this unit of  $F_1$  were absent and the remaining factors were organized most efficiently to cope with its absence. Thus marginal productivity theory is not best thought of as a theory of actual produced product, of those things whose causal pedigree includes the unit of the factor, but rather as a theory of the difference (subjunctively defined) made by the presence of a factor. If such a view were connected with justice, it would seem to fit best with an entitlement conception.

But if the notion of marginal product were so ineffective that factors' marginal products in actual situations of joint production could not be identified by hirers or purchasers of the factors, then the resulting distribution to factors would not be patterned in accordance with marginal product. Someone who viewed marginal productivity theory, where it was applicable, as a *patterned theory of justice*, might think that such situations of joint production and indeterminate marginal product provided an opportunity for some theory of justice to enter to determine appropriate exchange ratios. But an entitlement theorist would find acceptable whatever distribution resulted from the party's voluntary exchanges.\* The questions about the workability of marginal productivity theory are intricate ones.<sup>20</sup> Let us merely note here the strong personal incentive for owners of resources to converge to the marginal product, and the strong market pressures tending to produce this result. Employers of factors of productions are not all dolts who don't know what they're doing, transferring holdings they value to others on an irrational and arbitrary basis. Indeed, Rawls' position on inequalities requires that separate contributions to joint products be isolable, to some extent at least. For Rawls goes out of his way to argue that inequalities are justified if they serve to raise the position of the worst-off group in the society, if without the inequalities the worst-off group would be even more worse off. These serviceable inequalities stem, at least in part, from the necessity to provide incentives to certain people to perform various activities or fill various roles that not everyone can do equally well. (Rawls is *not* imagining that inequalities are needed to fill positions that everyone can do equally well, or that the most drudgery-filled positions that require the least skill will command the highest income.) But *to whom* are the incentives to be paid? To which performers of what activities? When it is necessary to provide incentives to some to perform their productive activities, there is no talk of a joint social product from which no individual's contribution can be disentangled. If the product was all that

\* Readers who believe that Marx's analysis of exchange relations between owners of capital and laborers undercuts the view that the set of holdings which results from voluntary exchange is legitimate, or who believe it a distortion to term such exchanges "voluntary," will find some relevant considerations aduced in Chapter 8.

inextricably joint, it couldn't be known that the extra incentives were going to the crucial persons; and it couldn't be known that the additional product produced by these now motivated people is greater than the expenditure to them in incentives. So it couldn't be known whether the provision of incentives was efficient or not, whether it involved a net gain or a net loss. But Rawls' discussion of justifiable inequalities presupposes that these things can be known. And so the claim we have imagined about the indivisible, nonpartitionable nature of the joint product is seen to dissolve, leaving the reasons for the view that social cooperation creates special problems of distributive justice otherwise not present, unclear if not mysterious.

#### TERMS OF COOPERATION AND THE DIFFERENCE PRINCIPLE

Another entry into the issue of the connection of social cooperation with distributive shares brings us to grips with Rawls' actual discussion. Rawls imagines rational, mutually disinterested individuals meeting in a certain situation, or abstracted from their other features not provided for in this situation. In this hypothetical situation of choice, which Rawls calls "the original position," they choose the first principles of a conception of justice that is to regulate all subsequent criticism and reform of their institutions. While making this choice, no one knows his place in society, his class position or social status, or his natural assets and abilities, his strength, intelligence, and so forth.

The principles of justice are chosen behind a veil of ignorance. This ensures that no one is advantaged or disadvantaged in the choice of principles by the outcome of natural chance or the contingency of social circumstances. Since all are similarly situated and no one is able to design principles to favor his particular condition, the principles of justice are the result of a fair agreement or bargain.<sup>21</sup>

What would persons in the original position agree to?

Persons in the initial situation would choose two . . . principles: the first requires equality in the assignment of basic rights and duties, while

the second holds that social and economic inequalities, for example, inequalities of wealth and authority are just only if they result in compensating benefits for everyone, and in particular for the least advantaged members of society. These principles rule out justifying institutions on the grounds that the hardships of some are offset by a greater good in the aggregate. It may be expedient but it is not just that some should have less in order that others may prosper. But there is no injustice in the greater benefits earned by a few provided that the situation of persons not so fortunate is thereby improved. The intuitive idea is that since everyone's well-being depends upon a scheme of cooperation without which no one could have a satisfactory life, the division of advantages should be such as to draw forth the willing cooperation of everyone taking part in it, including those less well situated. Yet this can be expected only if reasonable terms are proposed. The two principles mentioned seem to be a fair agreement on the basis of which those better endowed, or more fortunate in their social position, neither of which we can be said to deserve, could expect the willing cooperation of others when some workable scheme is a necessary condition of the welfare of all.<sup>22</sup>

This second principle, which Rawls specifies as the difference principle, holds that the institutional structure is to be so designed that the worst-off group under it is at least as well off as the worst-off group (not necessarily the same group) would be under any alternative institutional structure. If persons in the original position follow the minimax policy in making the significant choice of principles of justice, Rawls argues, they will choose the difference principle. Our concern here is not whether persons in the position Rawls describes actually would minimax and actually would choose the particular principles Rawls specifies. Still, we should question why individuals in the original position would choose a principle that focuses upon groups, rather than individuals. Won't application of the minimax principle lead each person in the original position to favor maximizing the position of the worst-off *individual*? To be sure, this principle would reduce questions of evaluating social institutions to the issue of how the unhappiest depressive fares. Yet avoiding this by moving the focus to groups (or representative individuals) seems *ad hoc*, and is inadequately motivated for those in the individual position.<sup>23</sup> Nor is it clear which groups are appropriately considered; why exclude the group of depressives or alcoholics or the representative paraplegic?

If the difference principle is not satisfied by some institutional

structure  $J$ , then under  $J$  some group  $G$  is worse off than it would be under another institutional structure  $I$  that satisfies the principle. If another group  $F$  is better off under  $J$  than it would be under the  $I$  favored by the difference principle, is this sufficient to say that under  $J$  "some . . . have less in order that others may prosper"? (Here one would have in mind that  $G$  has less in order that  $F$  prosper. Could one also make the same statement about  $I$ ? Does  $F$  have less under  $I$  in order that  $G$  may prosper?) Suppose that in a society the following situation prevailed:

1. Group  $G$  has amount  $A$  and group  $F$  has amount  $B$ , with  $B$  greater than  $A$ . Also things could be arranged differently so that  $G$  would have more than  $A$ , and  $F$  would have less than  $B$ . (The different arrangement might involve a mechanism to transfer some holdings from  $F$  to  $G$ .)

Is this sufficient to say

2.  $G$  is badly off because  $F$  is well off;  $G$  is badly off in order that  $F$  be well off;  $F$ 's being well off makes  $G$  badly off;  $G$  is badly off on account of  $F$ 's being well off;  $G$  is not better off because of how well off  $F$  is.

If so, does the truth of statement 2 depend on  $G$ 's being in a worse position than  $F$ ? There is yet another possible institutional structure  $K$  that transfers holdings from the worse-off group  $G$  to  $F$ , making  $G$  even more worse off. Does the possibility of  $K$  make it true to say that, under  $J$ ,  $F$  is not (even) better off because of how well off  $G$  is?

We do not normally hold that the truth of a subjunctive (as in 1) is alone sufficient for the truth of some indicative causal statement (as in 2). It would improve my life in various ways if you were to choose to become my devoted slave, supposing I could get over the initial discomfort. Is the cause of my present state your not becoming my slave? Because your enslaving yourself to a poorer person would improve his lot and worsen yours, are we to say that the poor person is badly off because you are as well off as you are; has he less in order that you may prosper? From

3. If  $P$  were to do act  $A$  then  $Q$  would not be in situation  $S$ .

we will conclude

4. *P*'s not doing *A* is responsible for *Q*'s being in situation *S*; *P*'s not doing *A* causes *Q* to be in *S*.

only if we *also* believe that

5. *P* ought to do act *A*, or *P* has a duty to do act *A*, or *P* has an obligation to do act *A*, and so forth.<sup>24</sup>

Thus the inference from 3 to 4, in this case, *presupposes* 5. One cannot argue from 3 to 4 as one step in order *to get to* 5. The statement that in a particular situation some have less in order that others may prosper is often based upon the very evaluation of a situation or an institutional framework that it is introduced to support. Since this evaluation does *not* follow merely from the subjunctive (for example, 1 or 3) an *independent* argument must be produced for it.\*

Rawls holds, as we have seen, that

since everyone's well-being depends upon a scheme of cooperation without which no one could have a satisfactory life, the division of advantages should be such as to draw forth the willing cooperation of everyone taking part in it, including those less well situated. Yet this can be expected only if reasonable terms are proposed. The two principles mentioned seem to be a fair agreement on the basis of which those better endowed or more fortunate in their social position . . . could expect the willing cooperation of others when some workable scheme is a necessary condition of the welfare of all.<sup>25</sup>

No doubt, the difference principle presents terms on the basis of which those less well endowed would be willing to cooperate. (What *better* terms could they propose for themselves?) But is this a fair agreement on the basis of which those *worse* endowed could expect the *willing* cooperation of others? With regard to the existence of gains from social cooperation, the situation is symmetrical. The better endowed gain by cooperating with the worse endowed, *and* the worse endowed gain by cooperating with the better endowed. Yet the difference principle is not neutral be-

\* Though Rawls does not clearly distinguish 2 from 1 and 4 from 3, I do not claim that he makes the illegitimate step of sliding from the latter subjunctive to the former indicative. Even so, the mistake is worth pointing out because it is an easy one to fall into, and it might appear to prop up positions we argue against.

tween the better and the worse endowed. Whence the asymmetry?

Perhaps the symmetry is upset if one asks *how much* each gains from the social cooperation. This question might be understood in two ways. How much do people benefit from social cooperation, as compared to their individual holdings in a *noncooperative* scheme? That is, how much is  $T_i - S_i$ , for each individual *i*? Or, alternatively, how much does each individual gain from general social cooperation, as compared, not with *no* cooperation, but with more limited cooperation? The latter is the more appropriate question with regard to general social cooperation. For failing general agreement on the principles to govern how the benefits of general social cooperation are to be held, not everyone will remain in a noncooperative situation if there is some other beneficial cooperative arrangement involving some, but not all, people, whose participants *can* agree. These people will participate in this more narrow cooperative arrangement. To focus upon the benefits of the better and the worse endowed cooperating together, we must try to imagine less extensive schemes of partitioned social cooperation in which the better endowed cooperate only among themselves and the worse endowed cooperate only among themselves, with no cross-cooperation. The members of both groups gain from the internal cooperation within their respective groups and have larger shares than they would if there were no social cooperation at all. An individual benefits from the wider system of extensive cooperation between the better and the worse endowed to the extent of his incremental gain from this wider cooperation; namely, the amount by which his share under a scheme of general cooperation is greater than it would be under one of limited intragroup (but not cross-group) cooperation. *General* cooperation will be of more benefit to the better or to the worse endowed if (to pick a simple criterion) the mean incremental gain from general cooperation (when compared with limited intragroup cooperation) is greater in one group than it is in the other.

One might speculate about whether there is an inequality between the groups' mean incremental gains and, if so, which way it goes. If the better-endowed group includes those who manage to accomplish something of great economic advantage to others, such as new inventions, new ideas about production or ways of doing

things, skill at economic tasks, and so on,\* it is difficult to avoid concluding that the *less* well endowed gain *more* than the better endowed do from the scheme of general cooperation. What follows from this conclusion? I do *not* mean to imply that the better endowed should get even more than they get under the entitlement system of general social cooperation.† What *does* follow from the conclusion is a deep suspicion of imposing, in the name of fairness, constraints upon voluntary social cooperation (and the set of

\* They needn't be *better endowed*, from birth. In the context in which Rawls uses it, all "better endowed" means is: accomplishes more of economic value, able to do this, has a high marginal product, and so forth. (The role that unpredictable factors play in this complicates imagining a prior partitioning of the two groups.) The text follows Rawls in categorizing persons as "better" and "worse" endowed only in order to criticize the considerations *he* adduces for his theory. The entitlement theory does not rest upon any assumption that the classification is an important one, or even a possible one, or upon any elitist presupposition.

Since the entitlement theorist does not accept the patterned principle "to each according to his natural endowment," he can easily grant that what an exercised endowment brings in the market will depend upon the endowments of others and how they choose to exercise them, upon the market-expressed desires of buyers, upon the alternate supply of what he offers and of what others may substitute for what he offers, and upon other circumstances summing the myriad choices and actions of others. Similarly, we saw earlier that the similar considerations Rawls adduces about the social factors upon which the marginal product of labor depends (*Theory of Justice*, p. 308) will not faze an entitlement theorist, even though they might undercut the rationale put forth by a proponent of the patterned principle of distribution according to marginal product.

† Supposing they could identify themselves and each other, they might *try* to exact a larger share by banding together as a group and bargaining jointly with the others. Given the large numbers of persons involved and the incentive for some of the better-endowed individuals to break ranks and reach separate agreements with the worse endowed, if such a coalition of the better endowed is unable to impose sanctions on its defectors it will dissolve. The better endowed remaining in the coalition may use boycott as a "sanction," and refuse to cooperate with a defector. To break the coalition, those less well endowed would have to (be able to) offer someone better endowed sufficient incentive to defect to make up for his loss through no longer being able to cooperate with the other better-endowed persons. Perhaps it would pay for someone to defect from the coalition only as part of a sizable group of defectors, which defecting group the initial coalition might try to keep small by special offers to individuals to defect *from it*, and so on. The problem is a complicated one, further complicated by the obvious fact (despite our use of Rawls' classificatory terminology) that there is no sharp line of cleavage between the endowments of people to determine which groups would form.

holdings that arises from it) so that those already benefiting most from this general cooperation benefit even more!

Rawls would have us imagine the worse-endowed persons say something like the following: "Look, better endowed: you gain by cooperating with us. If you want our cooperation you'll have to accept reasonable terms. We suggest these terms: We'll cooperate with you only if we get *as much as possible*. That is, the terms of our cooperation should give us that maximal share such that, if it was tried to give us more, we'd end up with less." How generous these proposed terms are might be seen by imagining that the better endowed make the almost symmetrical opposite proposal: "Look, worse endowed: you gain by cooperating with *us*. If you want our cooperation you'll have to accept reasonable terms. We propose these terms: We'll cooperate with you so long as *we* get as much as possible. That is, the terms of our cooperation should give us the maximal share such that, if it was tried to give us more, we'd end up with less." If these terms seem outrageous, as they are, why don't the terms proposed by those worse endowed seem the same? Why shouldn't the better endowed treat this latter proposal as beneath consideration, supposing someone to have the nerve explicitly to state it?

Rawls devotes much attention to explaining why those less well favored should not complain at receiving less. His explanation, simply put, is that because the inequality works for his advantage, someone less well favored shouldn't complain about it; he receives *more* in the unequal system than he would in an equal one. (Though he might receive still more in another unequal system that placed someone else below him.) But Rawls discusses the question of whether those *more* favored will or should find the terms satisfactory *only* in the following passage, where *A* and *B* are any two representative men with *A* being the more favored:

The difficulty is to show that *A* has no grounds for complaint. Perhaps he is required to have less than he might since his having more would result in some loss to *B*. Now what can be said to the more favored man? To begin with, it is clear that the well-being of each depends on a scheme of social cooperation without which no one could have a satisfactory life. Secondly, we can ask for the willing cooperation of everyone only if the terms of the scheme are reasonable. The difference principle, then, seems to be a fair basis on which those better endowed, or more



fortunate in their social circumstances, could expect others to collaborate with them when some workable arrangement is a necessary condition of the good of all.<sup>26</sup>

What Rawls imagines being said to the more favored men does *not* show that these men have no grounds for complaint, nor does it at all diminish the weight of whatever complaints they have. That the well-being of all depends on social cooperation without which no one could have a satisfactory life could also be said to the less well endowed by someone proposing any other principle, including that of maximizing the position of the best endowed. Similarly for the fact that we can ask for the willing cooperation of everyone only if the terms of the scheme are reasonable. The question is: What terms *would be* reasonable? What Rawls imagines being said thus far merely sets up his problem; it doesn't distinguish his proposed difference principle from the almost symmetrical counterproposal that we imagined the better endowed making, or from any other proposal. Thus, when Rawls continues, "The difference principle, then, seems to be a fair basis on which those best endowed, or more fortunate in their social circumstances, could expect others to collaborate with them when some workable arrangement is a necessary condition of the good of all," the presence of the "then" in his sentence is puzzling. Since the sentences which precede it are neutral between his proposal and any other proposal, the conclusion that the difference principle presents a fair basis for cooperation *cannot* follow from what precedes it in this passage. Rawls is merely repeating that it seems reasonable; hardly a convincing reply to anyone to whom it doesn't seem reasonable.\*

---

\* I treat Rawls' discussion here as one concerning better- and worse-endowed individuals who know they are so. Alternatively, one might imagine that *these* considerations are to be weighed by someone in the original position. ("If I turn out to be better endowed then . . . ; if I turn out to be worse endowed then . . .") But this construal will not do. Why would Rawls bother saying, "The two principles . . . seem to be a fair agreement on the basis of which those better endowed or more fortunate in their social position could expect the willing cooperation of others" (*Theory of Justice*, p. 15). Who is doing the expecting when? How is this to be translated into subjunctives to be contemplated by someone in the original position? Similarly, questions arise about Rawls' saying, "The difficulty is to show that *A* has no grounds for complaint. Perhaps he is required to have less than he might since his having more would result in

Rawls has not shown that the more favored man *A* has no grounds for complaint at being required to have less in order that another *B* might have more than he otherwise would. And he can't show this, since *A* *does* have grounds for complaint. Doesn't he?

---

some loss to *B*. *Now what can be said to the more favored man?* . . . The difference principle then seems to be a fair basis on which those better endowed . . . could expect others to collaborate with them . . ." (*Theory of Justice*, p. 103, my italics). Are we to understand this as: someone in the original position wonders what to say to himself as he then thinks of the possibility that he will turn out to be one of the better endowed? And does he then say that the difference principle *then* seems a fair basis for cooperation despite the fact that, and even while, he is contemplating the possibility that he is better endowed? Or does he say then that even later if and when he knows he is better endowed the difference principle will seem fair to him at that later time? And when are we to imagine him possibly complaining? Not while in the original position, for then he is agreeing to the difference principle. Nor does he worry, while in the process of deciding in the original position, that he will complain later. For he knows that he will have no cause to complain later at the effects of whatever principle he himself rationally will choose soon in the original position. Are we to imagine him complaining against himself? And isn't the answer to any later complaint, "You agreed to it (or you would have agreed to it if so originally positioned)"? What "difficulty" does Rawls concern himself with here? Trying to squeeze it into the original position makes it completely mysterious. And what is thinking of what is a "fair agreement" (sect. 3) or a "fair basis" (p. 103) doing here anyway, in the midst of the rational self-interested calculations of persons in the original position, who do not then knowingly possess, or at any rate utilize, particular moral notions?

I see no coherent way to incorporate how Rawls treats and speaks of the issue of the terms of cooperation between the better and the worse endowed into the structure and perspective of the original position. Therefore my discussion considers Rawls here as addressing himself to individuals *outside* the original position, either to better-endowed individuals or to his readers, to convince *them* that the difference principle which Rawls extracts from the original position is fair. It is instructive to compare how Rawls imagines justifying the social order to a person in the worst-off group in an unequal society. Rawls wants to tell this person that the inequalities work out to his advantage. This is told to someone who knows who he is: "The social order can be justified to everyone, and in particular to those who are least favored" (p. 103). Rawls does not want to say, "You would have gambled, and you lost," or any such thing, even "You chose it then in the original position"; nor does he wish merely to address someone in the original position. He also wants a consideration apart from the original position that will convince someone who knows of his inferior position in an unequal society. To say, "You have less in order that I may prosper," would *not* convince someone who knows of his inferior position, and Rawls rightly rejects it, even though its subjunctive analogue for someone in the original position, if we could make sense of this, would not be without force.

THE ORIGINAL POSITION AND  
END-RESULT PRINCIPLES

How can it have been supposed that these terms offered by the less well endowed are fair? Imagine a social pie somehow appearing so that *no one* has any claim at all on any portion of it, no one has any more of a claim than any other person; yet there must be unanimous agreement on how it is to be divided. Undoubtedly, apart from threats or holdouts in bargaining, an equal distribution would be suggested and found plausible as a solution. (It is, in Schelling's sense, a focal point solution.) If *somehow* the size of the pie wasn't fixed, and it was realized that pursuing an equal distribution somehow would lead to a smaller total pie than otherwise might occur, the people might well agree to an unequal distribution which raised the size of the least share. But in any actual situation, wouldn't this realization reveal something about differential claims on parts of the pie? Who is it that could make the pie larger, and would do it if given a larger share, but not if given an equal share under the scheme of equal distribution? To whom is an incentive to be provided to make this larger contribution? (There's no talk here of inextricably entangled joint product; it's known to *whom* incentives are to be offered, or at least to whom a bonus is to be paid after the fact.) Why doesn't this identifiable differential contribution lead to some differential entitlement?

If things fell from heaven like manna, and no one had any special entitlement to any portion of it, and no manna would fall unless all agreed to a particular distribution, and somehow the quantity varied depending on the distribution, then it is plausible to claim that persons placed so that they couldn't make threats, or hold out for specially large shares, would agree to the difference principle rule of distribution. But *is this* the appropriate model for thinking about how the things people produce are to be distributed? Why think the same results should obtain for situations where there *are* differential entitlements as for situations where there are not?

A procedure that founds principles of distributive justice on what rational persons who know nothing about themselves or their histories would agree to *guarantees that end-state principles of justice*

*will be taken as fundamental.* Perhaps some historical principles of justice are derivable from end-state principles, as the utilitarian tries to derive individual rights, prohibitions on punishing the innocent, and so forth, from *his* end-state principle; perhaps such arguments can be constructed even for the entitlement principle. But no historical principle, it seems, could be agreed to in the first instance by the participants in Rawls' original position. For people meeting together behind a veil of ignorance to decide who gets what, knowing nothing about any special entitlements people may have, will treat anything to be distributed as manna from heaven.\*

Suppose there were a group of students who have studied during a year, taken examinations, and received grades between 0 and 100 which they have not yet learned of. They are now gathered together, having no idea of the grade any one of them has received, and they are asked to allocate grades among themselves so that the grades total to a given sum (which is determined by the sum of the grades they actually have received from the teacher). First, let us suppose they are to decide jointly upon a particular distribution of grades; they are to give a particular grade to each identifiable one of them present at the meeting. Here, given sufficient restrictions on their ability to threaten each other, they probably would agree to each person receiving the same grade, to each person's grade being equal to the total divided by the number of people to be graded. Surely they would *not* chance upon the particular set of grades they already have received. Suppose next that there is posted on a bulletin board at their meeting a paper headed ENTITLEMENTS, which lists each person's name with a grade next to it, the listing being identical to the instructor's gradings. Still, this particular distribution will not be agreed to by those having done poorly. Even if they know what "entitlement" means (which perhaps we must suppose they don't, in order to match the

---

\* Do the people in the original position ever wonder whether *they* have the right to decide how everything is to be divided up? Perhaps they reason that since they are deciding this question, they must assume they are entitled to do so; and so particular people can't have particular entitlements to holdings (for then they wouldn't have the right to decide together on how all holdings are to be divided); and hence everything legitimately may be treated like manna from heaven.

absence of moral factors in the calculations of persons in Rawls' original position), why should they agree to the instructor's distribution? What self-interested reason to agree to it would they have?

Next suppose that they are unanimously to agree not to a *particular* distribution of grades, but rather to general principles to govern the distribution of grades. What principle would be selected? The equality principle, which gives each person the same grade, would have a prominent chance. And if it turned out that the total was variable depending upon how they divided it, depending on which of them got what grade, and a higher grade was desirable though they were not competing among each other (for example, each of them was competing for some position with the members of separate distinct groups), then the principle of distributing grades so as to maximize the lowest grades *might* seem a plausible one. Would these people agree to the non-end-state *historical* principle of distribution: give people grades according to how their examinations were evaluated by a qualified and impartial observer? \* If all the people deciding knew the particular distribution that would be yielded by this historical principle, they wouldn't agree to it. For the situation then would be equivalent to the earlier one of their deciding upon a particular distribution, in which we already have seen they would not agree to the entitlement distribution. Suppose then that the people do not know the particular distribution actually yielded by this historical principle. They cannot be led to select this historical principle because it looks just, or fair, to them; for no such notions are allowed to be at work in the original position. (Otherwise people would argue there, like here, about what justice requires.) Each person engages in a calculation to decide whether it will be in his own interests to accept this historical principle of distribution. Grades, under the historical prin-

\* I do not mean to assume that all teachers are such, nor even that learning in universities should be graded. All I need is some example of entitlement, the details of which the reader will have some familiarity with, to use to examine decision making in the original position. Grading is a simple example, though not a perfect one, entangled as it is with whatever ultimate social purposes the ongoing practice serves. We may ignore this complication, for their selecting the historical principle on the grounds that it effectively serves those purposes would illustrate our point below that their fundamental concerns and fundamental principles are end-state ones.

ciple, depend upon nature and developed intelligence, how hard the people have worked, accident, and so on, factors about which people in the original position know almost nothing. (It would be risky for someone to think that since he is reasoning so well in thinking about the principles, he must be one of the intellectually better endowed. Who knows what dazzling argument the others are reasoning their way through, and perhaps keeping quiet about for strategic reasons.) Each person in the original position will do something like assigning probability distributions to his place along these various dimensions. It seems unlikely that each person's probability calculations would lead to the historical-entitlement principle, in preference to every other principle. Consider the principle we may call the reverse-entitlement principle. It recommends drawing up a list of the historical entitlements in order of magnitude, and giving the most anyone is entitled to, to the person entitled to the least; the second most to the person entitled to the second least, and so on.<sup>27</sup> Any probability calculations of self-interested persons in Rawls' original position, or any probability calculations of the students we have considered, will lead them to view the entitlement and the reverse-entitlement principles as ranked equally insofar as their own self-interest is concerned! (What calculations could lead them to view one of the principles as superior to the other?) Their calculations will not lead them to select the entitlement principle.

The nature of the decision problem facing persons deciding upon principles in an original position behind a veil of ignorance limits them to end-state principles of distribution. The self-interested person evaluates any non-end-state principle on the basis of how it works out for him; his calculations about any principle focus on how he ends up under the principle. (These calculations include consideration of the labor he is yet to do, which does not appear in the grading example except as the sunk cost of the labor already done.) Thus for any principle, an occupant of the original position will focus on the distribution  $D$  of goods that it leads to, or a probability distribution over the distributions  $D_1, \dots, D_n$ ; it may lead to, and upon his probabilities of occupying each position in each  $D_i$  profile, supposing it to obtain. The point would remain the same if, rather than using personal probabilities, he uses some other decision rule of the sort discussed by decision theorists. In

these calculations, the only role played by the principle is that of generating a distribution of goods (or whatever else they care about) or of generating a probability distribution over distributions of goods. Different principles are compared solely by comparing the alternative distributions they generate. Thus the principles drop out of the picture, and each self-interested person makes a choice among alternative end-state distributions. People in the original position either directly agree to an end-state distribution or they agree to a principle; if they agree to a principle, they do it solely on the basis of considerations about end-state distributions. The *fundamental* principles they agree to, the ones they can all converge in agreeing upon, *must* be end-state principles.

Rawls' construction is incapable of yielding an entitlement or historical conception of distributive justice. The end-state principles of justice yielded by his procedure might be used in an attempt to *derive*, when conjoined with factual information, historical-entitlement principles, as derivative principles falling under a nonentitlement conception of justice.<sup>28</sup> It is difficult to see how such attempts could derive and account for the *particular* convolutions of historical-entitlement principles. And any derivations from end-state principles of approximations of the principles of acquisition, transfer, and rectification would strike one as similar to utilitarian contortions in trying to derive (approximations of) usual precepts of justice; they do not yield the particular result desired, and they produce the wrong reasons for the sort of result they try to get. If historical-entitlement principles are fundamental, then Rawls' construction will yield approximations of them at best; it will produce the wrong sorts of reasons for them, and its derived results sometimes will conflict with the precisely correct principles. The whole procedure of persons choosing principles in Rawls' original position presupposes that no historical-entitlement conception of justice is correct.

It might be objected to our argument that Rawls' procedure is designed to *establish* all facts about justice; there is no independent notion of entitlement, not provided by his theory, to stand on in criticizing his theory. But we do not need any *particular* developed historical-entitlement theory as a basis from which to criticize Rawls' construction. If *any* such fundamental historical-entitlement view is correct, then Rawls' theory is not. We are thus

able to make this structural criticism of the type of theory Rawls presents and the type of principles it must yield, without first having formulated fully a particular historical-entitlement theory as an alternative to his. We would be ill advised to accept Rawls' theory and his construal of the problem as one of which principles would be chosen by rational self-interested individuals behind a veil of ignorance, unless we were sure that no adequate historical-entitlement theory was to be gotten.

Since Rawls' construction doesn't yield a historical or entitlement conception of justice, there will be some feature(s) of his construction in virtue of which it doesn't. Have we done anything other than focus upon the particular feature(s), and say that this makes Rawls' construction incapable in principle of yielding an entitlement or historical conception of justice? This would be a criticism without any force at all, for in this sense we would have to say that the construction is incapable in principle of yielding any conception other than the one it actually yields. It seems clear that our criticism goes deeper than this (and I hope it is clear to the reader); but it is difficult to formulate the requisite criterion of depth. Lest this appear lame, let us add that as Rawls states the root idea underlying the veil of ignorance, that feature which is the most prominent in excluding agreement to an entitlement conception, it is to prevent someone from tailoring principles to his own advantage, from designing principles to favor his particular condition. But not only does the veil of ignorance do this; it ensures that no shadow of entitlement considerations will enter the rational calculations of ignorant, nonmoral individuals constrained to decide in a situation reflecting some formal conditions of morality.\* Perhaps, in a Rawls-like construction, some condition weaker than the veil of ignorance could serve to exclude the special tailoring of principles, or perhaps some other "structural-looking" feature of the choice situation could be formulated to mirror en-

---

\* Someone might think entitlement principles count as specially tailored in a morally objectionable way, and so he might reject my claim that the veil of ignorance accomplishes more than its stated purpose. Since to specially tailor principles is to tailor them *unfairly* for one's own advantage, and since the question of the fairness of the entitlement principle is precisely the issue, it is difficult to decide which begs the question: my criticism of the strength of the veil of ignorance, or the defense against this criticism which I imagine in this note.

entitlement considerations. But as it stands there is no reflection of entitlement considerations in any form in the situation of those in the original position; these considerations do not enter even to be overridden or outweighed or otherwise put aside. Since no glimmer of entitlement principles is built into the structure of the situation of persons in the original position, there is no way these principles could be selected; and Rawls' construction is incapable in principle of yielding them. This is not to say, of course, that the entitlement principle (or "the principle of natural liberty") couldn't be *written* on the list of principles to be considered by those in the original position. Rawls doesn't do even this, perhaps because it is so transparently clear that there would be no point in including it to be considered *there*.

#### MACRO AND MICRO

We noted earlier the objection which doubted whether there is any independent notion of entitlement. This connects with Rawls' insistence that the principles he formulates are to be applied only to the fundamental macrostructure of the whole society, and that no micro counterexample to them will be admissible. The difference principle is, on the face of it, *unfair* (though that will be of no concern to anyone deciding in the original position); and a wide gamut of counterexamples to it can be produced that focus on small situations that are easy to take in and manage. But Rawls does *not* claim the difference principle is to apply to every situation; only to the basic structure of the society. How are we to decide if it applies to that? Since we may have only weak confidence in our intuitions and judgments about the justice of the whole structure of society, we may attempt to aid our judgment by focusing on microsituations that we do have a firm grasp of. For many of us, an important part of the process of arriving at what Rawls calls "reflective equilibrium" will consist of thought experiments in which we try out principles in hypothetical microsituations. If, in our considered judgment, they don't apply there then they are not universally applicable. And we may think that since correct principles of justice *are* universally applicable, princi-

ples that fail for microsituations cannot be correct. Since Plato, at any rate, that has been our tradition; principles may be tried out in the large and in the small. Plato thought that writ large the principles are easier to discern; others may think the reverse.

Rawls, however, proceeds as though distinct principles apply to macro and micro contexts, to the basic structure of society and to the situations we can take in and understand. Are the fundamental principles of justice *emergent* in this fashion, applying only to the largest social structure yet not to its parts? Perhaps one thinks of the possibility that a whole social structure is just, even though none of its parts is, because the injustice in each part somehow balances out or counteracts another one, and the total injustice ends up being balanced out or nullified. But can a part satisfy the most fundamental principle of justice yet still clearly be unjust, apart from its failure to perform any supposed task of counterbalancing another existing injustice? Perhaps so, if a part involves some special domain. But surely a regular, ordinary, everyday part, possessing no very unusual features, should turn out to be just when it satisfies the fundamental principles of justice; otherwise, special explanations must be offered. One cannot say merely that one is speaking of principles to apply only to the fundamental structure, so that micro counterexamples do not tell. In virtue of what features of the basic structure, features not possessed by microcases, do special moral principles apply that would be unacceptable elsewhere?

There are special disadvantages to proceeding by focusing only on the intuitive justice of described complex wholes. For complex wholes are not easily scanned; we cannot easily keep track of everything that is relevant. The justice of a whole society may depend on its satisfying a number of distinct principles. These principles, though individually compelling (witness their application to a wide range of particular microcases), may yield surprising results when combined together. That is, one may be surprised at which, and only which, institutional forms satisfy all the principles. (Compare the surprise at discovering what, and only what, satisfies a number of distinct and individually compelling conditions of adequacy; and how illuminating such discoveries are.) Or perhaps it is one simple principle which is to be writ large, and what things look like when this is done is very surprising, at first. I am not

claiming that new *principles* emerge in the large, but that how the old microprinciples turn out to be satisfied in the large may surprise. If this is so, then one should not depend upon judgments about the whole as providing the only or even the major body of data against which to check one's principles. One major path to changing one's intuitive judgments about some complex whole is through seeing the larger and often surprising implications of principles solidly founded at the micro level. Similarly, discovering that one's judgments are wrong or mistaken often surely will involve overturning them by stringent applications of principles grounded on the micro level. For these reasons it is undesirable to attempt to protect principles by excluding microtests of them.

The only reason I have thought of for discounting microtests of the fundamental principles is that microsituations have particular entitlements built into them. Of course, continues the argument, the fundamental principles under consideration will run afoul of these entitlements, for the principles are to operate at a deeper level than such entitlements. Since they are to operate at the level that underlies such entitlements, no microsituation that includes entitlements can be introduced as an example by which to test these fundamental principles. Note that this reasoning grants that Rawls' procedure assumes that no fundamental entitlement view is correct, that it assumes there is some level so deep that no entitlements operate that far down.

May all entitlements be relegated to relatively superficial levels? For example, people's entitlements to the parts of their own bodies? An application of the principle of maximizing the position of those worst off might well involve forceable redistribution of bodily parts ("You've been sighted for all these years; now one—or even both—of your eyes is to be transplanted to others"), or killing some people early to use their bodies in order to provide material necessary to save the lives of those who otherwise would die young.<sup>29</sup> To bring up such cases is to sound slightly hysterical. But we are driven to such extreme examples in examining Rawls' prohibition on micro counterexamples. That not all entitlements in microcases are plausibly construed as superficial, and hence as illegitimate material by which to test out suggested principles, is made especially clear if we focus on those entitlements and rights that most clearly are not socially or institutionally

based. On what grounds are such cases, whose detailed specifications I leave to the ghoulish reader, ruled inadmissible? On what grounds can it be claimed that the fundamental principles of justice need apply only to the fundamental institutional structure of a society? (And couldn't we build such redistributive practices concerning bodily parts or the ending of people's lives into the fundamental structure of a society?)

It is ironic that we criticize Rawls' theory for its fundamental incompatibility with historical-entitlement conceptions of justice. For Rawls' theory itself describes a process (abstractly conceived) with a result. He does not present a direct deductive argument for his two principles of justice from other statements that entail them. Any deductive formulation of Rawls' argument would contain metastatements, statements about principles: such as, any principles agreed to by persons in a certain situation are correct. Combined with an argument showing that persons in that situation would agree to principles *P*, one can deduce that *P* is correct, and then deduce that *P*. At some places in the argument, "*P*" appears in quotes, distinguishing the argument from a direct deductive argument for the truth of *P*. Instead of a direct deductive argument, a situation and process are specified, and *any* principles that would emerge from that situation and process are held to constitute the principles of justice. (Here I ignore the complicated interplay between which principles of justice one wants to derive and which initial situation one specifies.) Just as for an entitlement theorist any set of holdings that emerges from a legitimate process (specified by the principle of transfer) is just, so for Rawls any set of principles that emerges from the original position by the constrained process of unanimous agreement is the set of (correct) principles of justice. Each theory specifies starting points and processes of transformation, and each accepts whatever comes out. According to each theory, whatever comes out is to be accepted because of its pedigree, its history. Any theory which gets to a process must start with something which is not *itself* justified by being the outcome of a process (otherwise, it should start farther back)—namely, either with general statements arguing for the fundamental priority of the process, or with the process itself. Entitlement theory and Rawls' theory each get to a process. Entitlement theory specifies a process for generating sets of holdings.

The three principles of justice (in acquisition, transfer, and rectification) that underlie this process, having this process as their subject matter, are themselves process principles rather than end-state principles of distributive justice. They specify an ongoing process, *without* fixing how it is to turn out, *without* providing some external patterned criterion it must meet. Rawls' theory arrives at a process *P* for generating principles of justice. This process *P* involves people in the original position agreeing to principles of justice behind a veil of ignorance. According to Rawls, any principles emerging from this process *P* will be the principles of justice. But this process *P* for generating principles of justice cannot, we already have argued, itself generate process principles as the fundamental principles of justice. *P* must generate end-state or end-result principles. Even though the difference principle, in Rawls' theory, is to apply to an ongoing and continuing institutional process (one that includes *derived* entitlements based upon institutional expectations under the principle, and derived elements of pure procedural justice, and so on), it is an end-result principle (but not a *current* time-slice principle). The difference principle fixes how the ongoing process is to turn out and provides an external patterned criterion it must meet; any process is rejected which fails to meet the test of the criterion. The mere fact that a principle regulates an ongoing institutional process does not make it a process principle. If it did, the utilitarian principle would also be a process principle, rather than the end-result principle it is.

The structure of Rawls' theory thus presents a dilemma. If processes are so great, Rawls' theory is defective because it is incapable of yielding process principles of justice. If processes are not so great, then insufficient support has been provided for the principles yielded by Rawls' process *P* for arriving at principles. Contract arguments embody the assumption that anything that emerges from a certain process is just. Upon the force of this fundamental assumption rests the force of a contract argument. Surely then no contract argument should be structured so as to preclude process principles being the fundamental principles of distributive justice by which to judge the institutions of a society; no contract argument should be structured so as to make it impossible that its results be of the same sort as the assumptions upon which it rests.<sup>30</sup> If processes are good enough to found a theory upon, they

are good enough to be the possible result of the theory. One can't have it both ways.

We should note that the difference principle is an especially strong kind of patterned end-state principle. Let us say that a principle of distribution is *organic* if an unjust distribution, according to the principle, can be gotten from one the principle deems just, by deleting (in imagination) some people and their distributive shares. Organic principles focus on features dependent upon the *overall* pattern. In contrast, patterned principles of the form "to each according to his score on a particular natural dimension *D*" are *not* organic principles. If a distribution satisfies this principle, it will continue to do so when some people and their holdings are deleted, for this deletion will not affect the ratios of the remaining people's holdings, or the ratios of their scores along the dimension *D*. These unchanged ratios will continue to be the same and will continue to satisfy the principle.

The difference principle *is* organic. If the least well-off group and their holdings are deleted from a situation, there is no guarantee that the resulting situation and distribution will maximize the position of the new least well-off group. Perhaps that new bottom group could have more if the top group had even less (though there was no way to transfer from the top group to the previous bottom group).\*

Failure to satisfy the deletion condition (that a distribution remains just under deletion of people and their holdings) marks off organic principles. Consider also the addition condition, which holds that if two distributions (over disjoint sets of individuals) are just then so is the distribution which consists of the combination of these two just distributions. (If the distribution on earth is just, and that on some planet of a distant star is just, then so is the sum distribution of the two.) Principles of distribution of the form "to each according to his score on natural dimension *D*" violate this condition, and therefore (let us say) are *nonaggregative*. For though within each group all ratios of shares match ratios of scores on

\* The difference principle thus creates *two* conflicts of interest: between those at the top and those at bottom; *and* between those in the middle and those at bottom, for if those at bottom were gone the difference principle might apply to improve the position of those in the middle, who would become the new bottom group whose position is to be maximized.

$D$ , they needn't match *between* the groups.\* The entitlement principle of justice in holdings satisfies both the deletion and the addition conditions; the entitlement principle is nonorganic and aggregative.

We should not leave the subject of the properties of the difference principle without mentioning the interesting but I think mistaken speculation of Thomas Scanlon that "there is no plausible principle which is distinct from the Difference Principle and intermediate between it and strict equality."<sup>31</sup> How can it be that no plausible egalitarian principle short of absolute equality would exclude great inequalities in order to achieve a *slight* benefit for the worst-off representative man? For the egalitarian, inequality is a cost, a minus-factor. The strict egalitarian doesn't allow any inequality at all, treating the cost of an inequality as infinite. The difference principle allows *any* amount of this cost provided there is *some* benefit (to the worst-off group) however small. This doesn't treat inequality as a significant cost. I have phrased my comments so that the following principle, call it Egalitarian General Principle 1, will leap to mind: An inequality is justified only if its benefits outweigh its costs. Following Rawls, suppose its benefits are only those to the worst-off group. How shall we measure its costs (and in a way so that they are comparable to its benefits)? The costs should represent the total amount of inequality in the society, which might be variously treated. So let us consider as the measure of inequality in a particular system (and hence its cost) the difference between the situation of the best-off representative man and the worst-off representative man. Let  $X_w$  be the share of the worst-off representative man under System  $X$ ; let  $X_B$  be the share of the best-off representative man under  $X$ . Let  $E$  be an efficient system of equality (in which everyone gets no less a share than in any other equal system). ( $E_B = E_w$ ) Thus we get the following First Specification of Egalitarian General Principle 1. (Other specifications would use other measures of inequality.) An un-

\* Let the second group have individuals who score half as much on  $D$  and have shares twice as large as the corresponding individuals in the first group, where in the first group the ratios between any two individuals' shares and their scores on  $D$  are the same. It follows that *within* the second group, the ratio of any two individuals' shares will be the same as the ratio of their scores. Yet between groups this identity of ratios will *not* hold.

equal system  $U$  is unjustified if  $U_B - U_w > U_w - E_w$ . (Or should it be  $\geq$ ?) An inequality is justified only if its benefit to the worst-off group ( $U_w - E_w$ ) is greater than (or equal to?) the cost of the inequality ( $U_B - U_w$ ). (Note that this involves measurement on an interval scale, and interpersonal comparisons.) This is an intermediate position the egalitarian might find attractive, and it is a stronger egalitarian principle than the difference principle.

There is an even *more* stringent egalitarian principle short of strict egalitarianism, supported by considerations similar to those which lead to the rejection of a simple cost-benefit principle for moral contexts.<sup>32</sup> This would give us Egalitarian General Principle 2: An unequal system  $U$  is justified only if a) its benefits outweigh its costs, *and* b) there is no *other* unequal system  $S$ , with lesser inequality, such that the *extra* benefits of  $U$  over  $S$  do not outweigh the extra costs of  $U$  over  $S$ . As before, treating  $X_B - X_w$  as the costs of the inequality in a system  $X$ , we get the following First Specification of Egalitarian General Principle 2: An unequal system  $U$  is justified only if:

$$a) U_w - E_w > U_B - U_w \text{ and}$$

b) There is no system  $S$  such that  $S_B - S_w < U_B - U_w$ , and

$$U_w - S_w \leq (U_B - U_w) - (S_B - S_w).$$

(Notice b) comes to: There is no system  $S$  with less inequality than  $U$ , such that the extra benefits of  $U$  over  $S$  are less than or equal to its extra costs.)

In *increasing* order of egalitarian stringency we have: the difference principle, the first specification of General Egalitarian Principle 1, the first specification of General Egalitarian Principle 2, and the principle of strict equality (choose  $E$ ). Surely an egalitarian would find the middle two more attractive than the difference principle. (Such an egalitarian might want to consider what changes in the structure of the Original Position or the nature of the persons in it, would lead to one of these egalitarian principles being chosen.) I do not myself, of course, suggest that these egalitarian principles are correct. But their consideration helps illuminate exactly how egalitarian the difference principle is, and make it implausible to claim it stands as the most egalitarian plausible principle short of strict equality. (However, perhaps



Scanlon means that any more stringent egalitarian principle would have to ascribe a cost to inequality, and no theoretical justification has been given which would enable one to ascribe a precise cost.)

There is one way we should mention whereby even more egalitarian principles might be gotten from Rawls' original position. Rawls imagines rational self-interested persons behind a veil of ignorance choosing principles to govern their institutions. He further imagines, in the third part of his book, that when raised in a society which embodies these principles, people thereby develop a sense of justice and a particular psychology (attitudes towards others, etc.). Call this Stage I of the argument. Stage II of the argument would involve taking *these* people who are the result of Stage I and the operation of a society in accordance with Stage I principles, and placing *them* in an original position. The Stage II original position contains individuals with the psychology and sense of justice which is the product of Stage I, rather than individuals who are (merely) rational and self-interested. Now these persons choose principles to govern the society they are to live in. Will the principles they choose in Stage II be the same principles chosen by the others in Stage I? If not, imagine people raised in a society embodying the Stage II principles, determine what psychology *they* would develop, and place *these* individuals, who are the products of Stage II, in a Stage III original position, and continue as before to iterate the process. We shall say that the iterated original position yields particular principles P if 1) there is a Stage  $n$  original position wherein P is chosen, and P is also chosen in the Stage  $n + 1$  original position, or 2) if new principles are chosen in each new stage of the original position, these principles converge to P at the limit. Otherwise, no particular principles are yielded by the iterated original position, e.g., succeeding stages of the original position oscillate between two sets of principles.

Are Rawls' two principles in fact yielded by the iterated original position, that is, at Stage II do the people with the psychology Rawls describes as resulting from the operation of his two principles of justice, themselves choose those very principles when *they* are placed in an original position? If so, this would strengthen Rawls' result. If not, we face the question of whether any principles are yielded by the original position; at what stage they are yielded (or are they yielded at the limit); and what precisely *those* princi-

ples are. This would seem to be an interesting area of investigation for those souls who choose to work, despite my arguments, within the Rawlsian framework.

#### NATURAL ASSETS AND ARBITRARINESS

Rawls comes closest to considering the entitlement system in his discussion of what he terms the system of natural liberty:

The system of natural liberty selects an efficient distribution roughly as follows. Let us suppose that we know from economic theory that under the standard assumptions defining a competitive market economy, income and wealth will be distributed in an efficient way, and that the particular efficient distribution which results in any period of time is determined by the initial distribution of assets, that is, by the initial distribution of income and wealth, and of natural talents and abilities. With each initial distribution, a definite efficient outcome is arrived at. Thus it turns out that if we are to accept the outcome as just, and not merely as efficient, we must accept the basis upon which over time the initial distribution of assets is determined.

In the system of natural liberty the initial distribution is regulated by the arrangements implicit in the conception of careers open to talents. These arrangements presuppose a background of equal liberty (as specified by the first principle) and a free market economy. They require a formal equality of opportunity in that all have at least the same legal rights of access to all advantaged social positions. But since there is no effort to preserve an equality or similarity, of social conditions, except insofar as this is necessary to preserve the requisite background institutions, the initial distribution of assets for any period of time is strongly influenced by natural and social contingencies. The existing distribution of income and wealth, say, is the cumulative effect of prior distributions of natural assets—that is, natural talents and abilities—as these have been developed or left unrealized, and their use favored or disfavored over time by social circumstances and such chance contingencies as accident and good fortune. Intuitively, the most obvious injustice of the system of natural liberty is that it permits distributive shares to be improperly influenced by these factors so arbitrary from a moral point of view.<sup>33</sup>

Here we have *Rawls'* reason for rejecting a system of natural liberty: it “permits” distributive shares to be improperly influenced by factors that are so arbitrary from a moral point of view. These

factors are: "prior distribution . . . of natural talents and abilities as these have been developed over time by social circumstances and such chance contingencies as accident and good fortune." Notice that there is no mention *at all* of how persons have chosen to develop their own natural assets. Why is that simply left out? Perhaps because such choices also are viewed as being the products of factors outside the person's control, and hence as "arbitrary from a moral point of view." "The assertion that a man deserves the superior character that enables him to make the effort to cultivate his abilities is equally problematic; for his character depends in large part upon fortunate family and social circumstances for which he can claim no credit."<sup>34</sup> (What view is presupposed here of character and its relation to action?) "The initial endowment of natural assets and the contingencies of their growth and nurture in early life are arbitrary from a moral point of view . . . the effort a person is willing to make is influenced by his natural abilities and skills and the alternatives open to him. The better endowed are more likely, other things equal, to strive conscientiously. . . ." <sup>35</sup> This line of argument can succeed in blocking the introduction of a person's autonomous choices and actions (and their results) only by attributing *everything* noteworthy about the person completely to certain sorts of "external" factors. So denigrating a person's autonomy and prime responsibility for his actions is a risky line to take for a theory that otherwise wishes to buttress the dignity and self-respect of autonomous beings; especially for a theory that founds so much (including a theory of the good) upon persons' choices. One doubts that the unexalted picture of human beings Rawls' theory presupposes and rests upon can be made to fit together with the view of human dignity it is designed to lead to and embody.

Before we investigate Rawls' reasons for rejecting the system of natural liberty, we should note the situation of those in the original position. The system of natural liberty is *one* interpretation of a principle that (according to Rawls) they *do* accept: social and economic inequalities are to be arranged so that they both are reasonably expected to be to everyone's advantage, and are attached to positions and offices open to all. It is left unclear whether the persons in the original position explicitly consider and choose among *all* the various interpretations of this principle, though this would

seem to be the most reasonable construal. (Rawls' chart on page 124 listing the conceptions of justice considered in the original position does *not* include the system of natural liberty.) Certainly they explicitly consider one interpretation, the difference principle. Rawls does not state why persons in the original position who considered the system of natural liberty would reject it. Their reason cannot be that it makes the resulting distribution depend upon a *morally* arbitrary distribution of natural assets. What we must suppose, as we have seen before, is that the self-interested calculation of persons in the original position does not (and cannot) lead them to adopt the entitlement principle. We, however, and Rawls, base our evaluations on different considerations.

Rawls has explicitly *designed* the original position and its choice situation so as to embody and realize his negative reflective evaluation of allowing shares in holdings to be affected by natural assets: "Once we decide to look for a conception of justice that nullifies the accidents of natural endowment and the contingencies of social circumstance. . . ." <sup>36</sup> (Rawls makes many scattered references to this theme of nullifying the accidents of natural endowment and the contingencies of social circumstance.) This quest crucially shapes Rawls' theory, and it underlies his delineation of the original position. It is not that persons who *did* deserve their natural endowments would choose differently if placed in Rawls' original position, but rather that, presumably, for such persons, Rawls would not hold that the principles of justice to govern *their* mutual relations were fixed by what they would choose in the original position. It is useful to remember how much of Rawls' construction rests upon this foundation. For example, Rawls argues that certain egalitarian demands are not motivated by envy but rather, because they are in accord with his two principles of justice, by resentment of injustice.<sup>37</sup> This argument can be undercut, as Rawls realizes,<sup>38</sup> if the very considerations which underlie the original position (yielding Rawls' two principles of justice) themselves embody or are based upon envy. So in addition to wanting to understand Rawls' rejection of alternative conceptions and to assess how powerful a criticism he makes of the entitlement conception, reasons internal to his theory provide motivation to explore the basis of the requirement that a conception of justice be geared to nullify differences in social circumstances and in natural

assets (and any differences in social circumstances they result in).

Why shouldn't holdings partially depend upon natural endowments? (They will also depend on how these are developed and on the uses to which they are put.) Rawls' reply is that these natural endowments and assets, being undeserved, are "arbitrary from a moral point of view." There are two ways to understand the relevance of this reply: It might be part of an argument to establish that the distributive effects of natural differences ought to be nullified, which I shall call the positive argument; or it might be part of an argument to rebut a possible counterargument holding that the distributive effects of natural differences oughtn't to be nullified, which I shall call the negative argument. Whereas the positive argument attempts to establish that the distributive effects of natural differences ought to be nullified, the negative one, by merely rebutting *one* argument that the differences oughtn't to be nullified, leaves open the possibility that (for other reasons) the differences oughtn't to be nullified. (The negative argument also leaves it possibly a matter of moral *indifference* whether the distributive effects of natural differences are to be nullified; note the difference between saying that something ought to be the case and saying that it's not that it oughtn't to be the case.)

#### THE POSITIVE ARGUMENT

We shall begin with the positive argument. How might the point that differences in natural endowments are arbitrary from a moral point of view function in an argument meant to establish that differences in holdings stemming from differences in natural assets ought to be nullified? We shall consider four possible arguments; the first, the following argument A:

1. Any person should morally deserve the holdings he has; it shouldn't be that persons have holdings they don't deserve.
2. People do not morally deserve their natural assets.
3. If a person's X partially determines his Y, and his X is undeserved then so is his Y.

Therefore,

4. People's holdings shouldn't be partially determined by their natural assets.

This argument will serve as a surrogate for other similar, more complicated ones.<sup>39</sup> But Rawls explicitly and emphatically *rejects* distribution according to moral desert.

There is a tendency for common sense to suppose that income and wealth, and the good things in life generally, should be distributed according to moral desert. Justice is happiness according to virtue. While it is recognized that this ideal can never be fully carried out, it is the appropriate conception [according to common sense] of distributive justice, at least as a *prima facie* principle, and society should try to realize it as circumstances permit. Now justice as fairness rejects this conception. Such a principle would not be chosen in the original position.<sup>40</sup>

Rawls could not, therefore, accept any premiss like the first premiss in argument A, and so no variant of this argument underlies his rejection of differences in distributive shares stemming from undeserved differences in natural assets. Not only does Rawls reject premiss 1, his theory is not coextensive with it. He favors giving incentives to persons if this most improves the lot of the least well off, and it often will be because of their natural assets that these persons will receive incentives and have larger shares. We noted earlier that the entitlement conception of justice in holdings, not being a patterned conception of justice, does not accept distribution in accordance with moral desert either. Any person may give to anyone else any holding he is entitled to, independently of whether the recipient morally deserves to be the recipient. To each according to the legitimate entitlements that legitimately have been transferred to him, is not a patterned principle.

If argument A and its first premiss are rejected, it is not obvious how to construct the positive argument. Consider next argument B:

1. Holdings ought to be distributed according to some pattern that is not arbitrary from a moral point of view.
2. That persons have different natural assets *is* arbitrary from a moral point of view.

Therefore,

3. Holdings ought not to be distributed according to natural assets.

But differences in natural assets might be *correlated* with other differences that are not arbitrary from a moral point of view and that are clearly of some possible moral relevance to distributional questions. For example, Hayek argued that under capitalism distribution generally is in accordance with perceived service to others. Since differences in natural assets will produce differences in ability to serve others, there will be some correlation of differences in distribution with differences in natural assets. The principle of the system is *not* distribution in accordance with natural assets; but differences in natural assets will lead to differences in holdings under a system whose principle is distribution according to perceived service to others. If conclusion 3 above is to be interpreted in extension so as to exclude this, it should be made explicit. But to add the premiss that any pattern that has some roughly coextensive description that is arbitrary from a moral point of view is itself arbitrary from a moral point of view would be far too strong, because it would yield the result that *every* pattern is arbitrary from a moral point of view. Perhaps the crucial thing to be avoided is not mere coextensiveness, but rather some morally arbitrary feature's *giving rise to* differences in distributive shares. Thus consider argument C:

1. Holdings ought to be distributed according to some pattern that is not arbitrary from a moral point of view.
2. That persons have different natural assets is arbitrary from a moral point of view.
3. If part of the explanation of why a pattern contains differences in holdings is that other differences in persons give rise to these differences in holdings, and if these other differences are arbitrary from a moral point of view, then the pattern also is arbitrary from a moral point of view.

Therefore,

4. Differences in natural assets should not give rise to differences in holdings among persons.

Premiss 3 of this argument holds that any moral arbitrariness that underlies a pattern infects the pattern and makes it too morally arbitrary. But any pattern will have some morally arbitrary facts as part of the explanation of how it arises, including the pattern

proposed by Rawls. The difference principle operates to give some persons larger distributive shares than others; which persons receive these larger shares will depend, at least partially, on differences between these persons and others, differences that are arbitrary from a moral point of view, for some persons with special natural assets will be offered larger shares as an incentive to use these assets in certain ways. Perhaps some premiss similar to 3 can be formulated so as to exclude what Rawls wishes to exclude while not excluding his *own* view. Still, the resulting argument would *assume* that the set of holdings should realize some pattern.

Why should the set of holdings be patterned? Patterning is *not* intrinsic to a theory of justice, as we have seen in our presentation of the entitlement theory: a theory that focuses upon the underlying principles that generate sets of holdings rather than upon the pattern a set of holdings realizes. If it be denied that the theory of these underlying principles *is* a separate theory of distributive justice, rather than merely a collection of diverse considerations from other areas, then the question becomes one of whether there *is* any separate subject of distributive justice which requires a separate theory.

On the manna-from-heaven model given earlier, there might be a more compelling reason to search for a pattern. But since things come into being already held (or with agreements already made about how they are to be held), there is no need to search for some pattern for unheld holdings to fit; and since the process whereby holdings actually come into being or are shaped, itself needn't realize any particular pattern, there is no reason to expect any pattern to result. The situation is not an appropriate one for wondering, "After all, what is to become of these things; what are we to do with them." In the non-manna-from-heaven world in which things have to be made or produced or transformed by people, there is no separate process of distribution for a theory of distribution to be a theory of. The reader will recall our earlier argument that (roughly) any set of holdings realizing a particular pattern may be transformed by the voluntary exchanges, gifts, and so forth, of the persons having the holdings under the pattern into *another* set of holdings that does not fit the pattern. The view that holdings *must* be patterned perhaps will seem less plausible when it is seen to have the consequence that people may not choose to

do acts that upset the patterning, even with things they legitimately hold.

There is another route to a patterned conception of justice that, perhaps, should be mentioned. Suppose that each morally legitimate fact has a "unified" explanation that shows it is morally legitimate, and that *conjunctions* fall into the domain of facts to be explained as morally legitimate. If  $p$ , and  $q$  are each morally legitimate facts, with their respective explanations as morally legitimate being  $P$ , and  $Q$ , then if  $p \wedge q$  is also to be explained as morally legitimate, and if  $P \wedge Q$  does not constitute a "unified" explanation (but is a mere conjunction of different explanations), then some further explanation will be needed. Applying this to holdings, suppose there are separate entitlement explanations showing the legitimacy of my having my holdings, and of your having yours, and the following question is asked: "Why is it legitimate that I hold what I do *and* you hold what you do; why is that joint fact *and all the relations contained within it* legitimate?" If the conjunction of the two separate explanations will not be held to explain in a unified manner the legitimacy of the joint fact (whose legitimacy is not viewed as being constituted by the legitimacy of its constituent parts), then some patterned principles of distribution would appear to be necessary to show its legitimacy, and to legitimate any nonunit set of holdings.

With scientific explanation of particular facts, the usual practice is to consider some conjunctions of explained facts as not requiring separate explanation, but as being explained by the conjunctions of the explanations of the conjuncts. (If  $E_1$  explains  $e_1$  and  $E_2$  explains  $e_2$  then  $E_1 \wedge E_2$  explains  $e_1 \wedge e_2$ .) If we required that any two conjuncts and any  $n$ -place conjunction had to be explained in some unified fashion, and not merely by the conjunction of separate and disparate explanations, then we would be driven to reject most of the usual explanations and to search for an underlying pattern to explain what appear to be separate facts. (Scientists, of course, often do offer a unified explanation of apparently separate facts.) It would be well worth exploring the interesting consequences of refusing to treat, even in the first instance, any two facts as legitimately separable, as having separate explanations whose conjunction is all there is to the explanation of them. What would our theories of the world look like if we required unified explanations

of *all* conjunctions? Perhaps an extrapolation of how the world looks to paranoid persons. Or, to put it undisparagingly, the way it appears to persons having certain sorts of dope experiences. (For example, the way it sometimes appears to me after smoking marijuana.) Such a vision of the world differs fundamentally from the way we normally look at it; it is surprising at first that a simple condition on the adequacy of explanations of conjunctions leads to it, until we realize that such a condition of adequacy must lead to a view of the world as deeply and wholly patterned.

A similar condition of adequacy on explanations of the moral legitimacy of conjunctions of separate morally legitimate facts would lead to a view that requires sets of holdings to exhibit an overall patterning. It seems unlikely that there will be compelling arguments for imposing such a principle of adequacy. Some may find such a unified vision plausible for only one realm; for example, in the moral realm concerning sets of holdings, but not in the realm of ordinary nonmoral explanation, or vice versa. For the case of explaining nonmoral facts, the challenge would be to produce such a unified theory. Were one produced that introduced novel considerations and explained no *new* facts (other than conjunctions of old ones) the decision as to its acceptability might be a difficult one and would depend largely on how explanatorily satisfying was the new way we saw the old facts. In the case of moral explanations and accounts which show the moral legitimacy of various facts, the situation is somewhat different. First, there is even less reason (I believe) to suppose a unified explanation appropriate and necessary. There is less need for a *greater* degree of explanatory unity than that provided when the same underlying principles for generating holdings appear in different explanations. (Rawls' theory, which contains elements of what he calls pure procedural justice, does not satisfy a strong condition of adequacy for explaining conjunctions and entails that such a condition cannot be satisfied.) Secondly, there is more danger than in the scientific case that the demand for a unified explanation will shape the "moral facts" to be explained. ("It can't be that both of those *are* facts for there's no unified patterned explanation that would yield them both.") Hence success in finding a unified explanation of such seriously primed facts will leave it unclear how well supported the explanatory theory is.

I turn now to our final positive argument which purports to derive the conclusion that distributive shares shouldn't depend upon natural assets from the statement that the distribution of natural assets is morally arbitrary. This argument focuses on the notion of equality. Since a large part of Rawls' argument serves to justify or show acceptable a particular deviation from equal shares (some may have more if this serves to improve the position of those worst off), perhaps a reconstruction of his underlying argument that places equality at its center will be illuminating. Differences between persons (the argument runs) are arbitrary from a moral point of view if there is no moral argument for the conclusion that there ought to be the differences. Not all such differences will be morally objectionable. That there is no such moral argument will seem important only in the case of those differences we believe oughtn't to obtain unless there is a moral reason establishing that they ought to obtain. There is, so to speak, a presumption against certain differences that can be overridden (can it merely be neutralized?) by moral reasons; in the absence of any such moral reasons of sufficient weight, there ought to be equality. Thus we have argument D:

1. Holdings ought to be equal, unless there is a (weighty) moral reason why they ought to be unequal.
2. People do not deserve the ways in which they differ from other persons in natural assets; there is no moral reason why people ought to differ in natural assets.
3. If there is no moral reason why people differ in certain traits, then their actually differing in these traits does not provide, and cannot give rise to, a moral reason why they should differ in other traits (for example, in holdings).

Therefore,

4. People's differing in natural assets is not a reason why holdings ought to be unequal.
5. People's holdings ought to be equal unless there is some other moral reason (such as, for example, raising the position of those worst off) why their holdings ought to be unequal.

Statements similar to the third premiss will occupy us shortly. Here let us focus on the first premiss, the equality premiss. Why ought people's holdings to be equal, in the absence of special moral reason to deviate from equality? (Why think there *ought* to

be *any* particular pattern in holdings?) Why is equality the rest (or rectilinear motion) position of the system, deviation from which may be caused only by moral forces? Many "arguments" for equality merely *assert* that differences between persons are arbitrary and must be justified. Often writers state a presumption in favor of equality in a form such as the following: "Differences in treatment of persons need to be justified."<sup>41</sup> The most favored situation for this sort of assumption is one in which there is one person (or group) treating everyone, a person (or group) having *no* right or entitlement to bestow the particular treatment as they wish or even whim. But if I go to one movie theater rather than to another adjacent to it, need I justify my different treatment of the two theater owners? Isn't it enough that I felt like going to one of them? That differences in treatment need to be justified *does* fit contemporary *governments*. Here there is a centralized process treating all, with no entitlement to bestow treatment according to whim. The major portion of distribution in a free society does not, however, come through the actions of the government, nor does failure to overturn the results of the localized individual exchanges constitute "state action." When there is no *one* doing the treating, and all are entitled to bestow their holdings as they wish, it is not clear why the maxim that differences in treatment must be justified should be thought to have extensive application. Why must differences between persons be justified? Why think that we must change, or remedy, or compensate for any inequality which can be changed, remedied, or compensated for? Perhaps here is where social cooperation enters in: though there is no presumption of equality (in, say, primary goods, or things people care about) among all persons, perhaps there is one among persons cooperating together. But it is difficult to see an argument for this; surely not all persons who cooperate together explicitly agree to this presumption as one of the terms of their mutual cooperation. And its acceptance would provide an unfortunate incentive for well-off persons to refuse to cooperate with, or to allow any of their number to cooperate with, some distant people who are less well off than any among them. For entering into such social cooperation, beneficial to those less well off, would seriously worsen the position of the well-off group by creating relations of presumptive equality between themselves and the worse-off group. In the next chapter I

shall consider the major recent argument for equality, one which turns out to be unsuccessful. Here we need only note that the connection argument D forges between not deserving natural assets and some conclusion about distributive shares *assumes* equality as a norm (that can be deviated from with, and only with, moral reason); and hence argument D itself cannot be used to establish any such conclusion about equality.

#### THE NEGATIVE ARGUMENT

Unsuccessful in our quest for a convincing positive argument to connect the claim that people don't deserve their natural assets with the conclusion that differences in holdings ought not to be based upon differences in natural assets, we now turn to what we called the negative argument: the use of the claim that people don't deserve their natural assets to rebut a possible counterargument to Rawls' view. (If the equality argument D were acceptable, the negative task of rebutting possible counterconsiderations would form part of the positive task of showing that a presumption for equality holds unoverridden in a particular case.) Consider the following possible counterargument E to Rawls:

1. People deserve their natural assets.
2. If people deserve X, they deserve any Y that flows from X.
3. People's holdings flow from their natural assets.

Therefore,

4. People deserve their holdings.
5. If people deserve something, then they ought to have it (and this overrides any presumption of equality there may be about that thing).

Rawls would rebut this counterargument to his position by denying its first premiss. And so we see *some* connection between the claim that the distribution of natural assets is arbitrary and the statement that distributive shares should not depend upon natural assets. However, no great weight can be placed upon *this* connection. For there are other counterarguments, in a similar vein; for example the argument F that begins:

1. If people have X, and their having X (whether or not they deserve to have it) does *not* violate anyone else's (Lockean) right or entitlement to X, and Y flows from (arises out of, and so on) X by a process that does not itself violate anyone's (Lockean) rights or entitlements,\* then the person is entitled to Y.
2. People's having the natural assets they do does not violate anyone else's (Lockean) entitlements or rights.

and goes on to argue that people are entitled to what they make, to the products of their labor, to what others give them or exchange. It is not true, for example, that a person earns Y (a right to keep a painting he's made, praise for writing *A Theory of Justice*, and so on) only if he's earned (or otherwise *deserves*) whatever he used (including natural assets) in the process of earning Y. Some of the things he uses he just may *have*, not illegitimately. It needn't be that the foundations underlying desert are themselves deserved, *all the way down*.

At the very least, we can parallel these statements about desert with ones about entitlements. And if, correctly, we describe people as entitled to their natural assets even if it's not the case that they can be said to deserve them, then the argument parallel to E above, with "are entitled to" replacing "deserve" throughout, *will* go through. This gives us the acceptable argument G:

1. People are entitled to their natural assets.
2. If people are entitled to something, they are entitled to whatever flows from it (via specified types of processes).
3. People's holdings flow from their natural assets.

Therefore,

4. People are entitled to their holdings.
5. If people are entitled to something, then they ought to have it (and

---

\* A process, we might strengthen the antecedent by adding, of the sort that would create an entitlement to Y if the person were entitled to X. I use "Lockean" rights and entitlements to refer to those (discussed in Part I) against force, fraud, and so on, which are to be recognized in the minimal state. Since I believe these are the only rights and entitlements people possess (apart from those they specially acquire), I needn't have included the specification to Lockean rights. One who believes some have a right to the fruits of others' labor will deny the truth of the first premiss as stated. If the Lockean specification were not included, he might grant the truth of 1, while denying that of 2 or of later steps.

this overrides any presumption of equality there may be about holdings).

Whether or not people's natural assets are arbitrary from a moral point of view, they are entitled to them, and to what flows from them.\*

A recognition of people's entitlements to their natural assets (the first premiss of argument G) might be necessary to avoid the stringent application of the difference principle which would lead, we already have seen, to even stronger property rights in other persons than redistributive theories usually yield. Rawls feels that he avoids this<sup>42</sup> because people in his original position rank the principle of liberty as lexicographically prior to the difference principle, applied not only to economic well-being but to health, length of life, and so on. (However, see note 29 above.)

We have found no cogent argument to (help) establish that differences in holding arising from differences in natural assets should be eliminated or minimized. Can the theme that people's natural assets are arbitrary from a moral point of view be used differently, for example, to justify a certain *shaping* of the original position? Clearly if the shaping is designed to nullify differences in holdings due to differences in natural assets, we need an argument for this goal, and we are back to our unsuccessful quest for the route to the conclusion that such differences in holdings ought to be nullified. Instead, the shaping might take place by excluding the participants in the original position from knowing of their own natural endowments. In this way the fact that natural endowments are arbitrary from a moral point of view would help to impose and to justify the veil of ignorance. But how does it do this; why should knowledge of natural endowments be excluded from the original

\* If nothing of moral significance could flow from what was arbitrary, then no particular person's existence could be of moral significance, since which of the many sperm cells succeeds in fertilizing the egg cell is (so far as we know) arbitrary from a moral point of view. This suggests another, more vague, remark directed to the spirit of Rawls' position rather than to its letter. Each existing person is the product of a process wherein the one sperm cell which succeeds is no more deserving than the millions that fail. Should we wish that process had been "fairer" as judged by Rawls' standards, that all "inequities" in it had been rectified? We should be apprehensive about any principle that would condemn morally the very sort of process that brought us to be, a principle that therefore would undercut the legitimacy of our very existing.

position? Presumably the underlying principle would be that if any particular features are arbitrary from a moral point of view, then persons in the original position should not know they possess them. But this would exclude their knowing *anything* about themselves, for each of their features (including rationality, the ability to make choices, having a life span of more than three days, having a memory, being able to communicate with other organisms like themselves) will be based upon the fact that the sperm and ovum which produced them contained particular genetic material. The physical fact that those particular gametes contained particular organized chemicals (the genes for people rather than for muskrats or trees) is arbitrary *from a moral point of view*; it is, from a moral point of view, an accident. Yet the persons in the original position are to know some of their attributes.

Perhaps we are too quick when we suggest excluding knowledge of rationality, and so forth, merely because these features *arise from* morally arbitrary facts. For these features also have moral significance; that is, moral facts depend upon or arise from them. Here we see an ambiguity in saying that a fact is arbitrary from a moral point of view. It might mean that there is no moral reason why the fact ought to be that way, or it might mean that the fact's being that way is of no moral significance and has no moral consequences. Rationality, the ability to make choices, and so on, are not morally arbitrary in this second sense. But if they escape exclusion on this ground, now the problem is that the natural assets, knowledge of which Rawls wishes to exclude from the original position, are not morally arbitrary in this sense either. At any rate, the entitlement theory's claim that moral entitlements may arise from or be partially based upon such facts is what is now at issue. Thus, in the absence of an argument to the effect that differences in holdings due to differences in natural assets ought to be nullified, it is not clear how anything about the original position can be based upon the (ambiguous) claim that differences in natural assets are arbitrary from a moral point of view.



## COLLECTIVE ASSETS

Rawls' view seems to be that everyone has some entitlement or claim on the totality of natural assets (viewed as a pool), with no one having differential claims. The distribution of natural abilities is viewed as a "collective asset."<sup>43</sup>

We see then that the difference principle represents, in effect, an agreement to regard the distribution of natural talents as a common asset and to share in the benefits of this distribution whatever it turns out to be. Those who have been favored by nature, whoever they are, may gain from their good fortune only on terms that improve the situation of those who have lost out. . . . No one deserves his greater natural capacity nor merits a more favorable starting place in society. But it does not follow that one should eliminate these distinctions. There is another way to deal with them. The basic structure can be arranged so that these contingencies work for the good of the least fortunate.<sup>44</sup>

People will differ in how they view regarding natural talents as a common asset. Some will complain, echoing Rawls against utilitarianism,<sup>45</sup> that this "does not take seriously the distinction between persons"; and they will wonder whether any reconstruction of Kant that treats people's abilities and talents as resources for others can be adequate. "The two principles of justice . . . rule out even the tendency to regard men as means to one another's welfare."<sup>46</sup> Only if one presses *very* hard on the distinction between men and their talents, assets, abilities, and special traits. Whether any coherent conception of a person remains when the distinction is so pressed is an open question. Why we, thick with particular traits, should be cheered that (only) the thus purified men within us are not regarded as means is also unclear.

People's talents and abilities *are* an asset to a free community; others in the community benefit from their presence and are better off because they are there rather than elsewhere or nowhere. (Otherwise they wouldn't choose to deal with them.) Life, over time, is not a constant-sum game, wherein if greater ability or effort leads to some getting more, that means that others must lose. In a free society, people's talents do benefit others, and not only themselves. Is it the extraction of even more benefit to others that is

supposed to justify treating people's natural assets as a collective resource? What justifies this extraction?

No one deserves his greater natural capacity nor merits a more favorable starting place in society. But it does not follow that one should eliminate these distinctions. There is another way to deal with them. The basic structure can be arranged so that these contingencies work for the good of the least fortunate.<sup>47</sup>

And if there weren't "another way to deal with them"? Would it then follow that one should eliminate these distinctions? What exactly would be contemplated in the case of natural assets? If people's assets and talents *couldn't* be harnessed to serve others, would something be done to remove these exceptional assets and talents, or to forbid them from being exercised for the person's own benefit or that of someone else he chose, even though this limitation wouldn't improve the absolute position of those somehow unable to harness the talents and abilities of others for their own benefit? Is it so implausible to claim that envy underlies this conception of justice, forming part of its root notion?\*

---

\* Will the lexicographic priority that Rawls claims for liberty in the original position prevent the difference principle from requiring a head tax on assets and abilities? The legitimacy of a head tax is *suggested* by Rawls' speaking of "collective assets" and "common assets." Those underutilizing their assets and abilities are misusing a public asset. (Squandering public property?) Rawls may intend no such strong inferences from his terminology, but we need to hear more about why those in the original position wouldn't accept the strong interpretation. The notion of liberty needs elaboration which is to exclude a head tax and yet allow the other taxation schemes. Assets and abilities can be harnessed without a head tax; and "harnessing" is an appropriate term—as it would be for a horse harnessed to a wagon which doesn't *have* to move ever, but if it does, it must draw the wagon along.

With regard to envy, the difference principle, applied to the choice between either *A* having ten and *B* having five or *A* having eight and *B* having five, would favor the latter. Thus, despite Rawls' view (pp. 79–80), the difference principle is inefficient in that it sometimes will favor a status quo against a Pareto-better but more unequal distribution. The inefficiency could be removed by shifting from the simple difference principle to a staggered difference principle, which recommends the maximization of the position of the least well-off group, and *subject to that constraint* the maximization of the position of the next least well-off group, and this point also is made by A. K. Sen (*Collective Choice and Social Welfare*, p. 138, note) and is acknowledged by Rawls (p. 83). But such a staggered principle does not embody a presumption in favor of equality

We have used our entitlement conception of justice in holdings to probe Rawls' theory, sharpening our understanding of what the entitlement conception involves by bringing it to bear upon an alternative conception of distributive justice, one that is deep and elegant. Also, I believe, we have probed deep-lying inadequacies in Rawls' theory. I am mindful of Rawls' reiterated point that a theory cannot be evaluated by focusing upon a single feature or part of it; instead the whole theory must be assessed (the reader will not know how whole a theory can be until he has read all of Rawls' book), and a perfect theory is not to be expected. However we have examined an important part of Rawls' theory, and its crucial underlying assumptions. I am as well aware as anyone of how sketchy my discussion of the entitlement conception of justice in holdings has been. But I no more believe we need to have formulated a complete alternative theory in order to reject Rawls' undeniably great advance over utilitarianism, than Rawls needed a complete alternative theory before he could reject utilitarianism. What more does one need or can one have, in order to begin progressing toward a better theory, than a sketch of a plausible alternative view, which from its very different perspective highlights the inadequacies of the best existing well-worked-out theory? Here, as in so many things, we learn from Rawls.

We began this chapter's investigation of distributive justice in order to consider the claim that a state more extensive than the minimal state could be justified on the grounds that it was necessary, or the most appropriate instrument, to achieve distributive justice. According to the entitlement conception of justice in holdings that we have presented, there is no argument based upon the first two principles of distributive justice, the principles of acquisition and of transfer, for such a more extensive state. If the set of holdings is properly generated, there is no argument for a more extensive state based upon distributive justice.<sup>48</sup> (Nor, we have claimed, will the Lockean proviso actually provide occasion for a more extensive state.) If, however, these principles are violated, the principle of rectification comes into play. Perhaps it is

---

of the sort used by Rawls. How then could Rawls justify an inequality *special* to the staggered principle to someone in the least well-off group? Perhaps these issues underlie the unclarity (see p. 83) as to whether Rawls accepts the staggered principle.

best to view some patterned principles of distributive justice as rough rules of thumb meant to approximate the general results of applying the principle of rectification of injustice. For example, lacking much historical information, and assuming (1) that victims of injustice generally do worse than they otherwise would and (2) that those from the least well-off group in the society have the highest probabilities of being the (descendants of) victims of the most serious injustice who are owed compensation by those who benefited from the injustices (assumed to be those better off, though sometimes the perpetrators will be others in the worst-off group), then a *rough* rule of thumb for rectifying injustices might seem to be the following: organize society so as to maximize the position of whatever group ends up least well-off in the society. This particular example may well be implausible, but an important question for each society will be the following: given *its* particular history, what operable rule of thumb best approximates the results of a detailed application in that society of the principle of rectification? These issues are very complex and are best left to a full treatment of the principle of rectification. In the absence of such a treatment applied to a particular society, one *cannot* use the analysis and theory presented here to condemn any particular scheme of transfer payments, unless it is clear that no considerations of rectification of injustice could apply to justify it. Although to introduce socialism as the punishment for our sins would be to go too far, past injustices might be so great as to make necessary in the short run a more extensive state in order to rectify them.