Chapter 8. Resource Egalitarianism

As the discussion in the previous chapter has shown, bargaining theory does not succeed in producing an uncontroversial specification of our conception of fairness. Rival bargaining solutions, employing different axiomatic constraints, suggest different ways of resolving even very simple distribution problems. It is unclear how one would decide which solution is best. Each appears to have its own merits, along with its own drawbacks.

On top of this, bargaining theory is vulnerable to an even more fundamental set of objections. All of the solutions presented in the previous chapter agree in the presumption that the correct object of equalization – the equalisandum – is welfare, understood here as utility, or the satisfaction of personal preference. Some philosophers have argued, however, that this is simply the wrong thing to be concerned with in a theory of distributive justice. These philosophers draw a distinction between welfare, either anticipated or actual, and the resources that are required to satisfy these preferences. They then ask which it is the responsibility of society to equalize. When we agree to cooperate, do we agree that the ultimate outcome should be one that all individuals find equally satisfactory, or do we simply agree that more tangible fruits of cooperation – the goods and services that we produce – should be distributed out equally?

The distinction between welfare and resources could be safely ignored in the discussion of efficiency, because the efficiency principle deals only with win-win transformations. Giving people more resources will in general lead to increased satisfaction of their actual preferences, and so it is not necessarily important to distinguish between the two. However, once we start talking about redistribution, we are talking about win-lose transformations. When we arrange transfers between individuals, we need to state with much greater precision exactly what is to be transferred, and how much of it is to be transferred. Thus the question of the appropriate equalisandum for the theory of justice becomes far more pressing.

To see how this issue can be important, consider the difference between equalizing expected utility and equalizing payoffs. The Nash bargaining solution equalizes expected utility. In Nash's toy-division problem, the bargaining solution may not assign a specific toy to one or the other boy, but may use a lottery that gives each one some specific chance of getting it. If one of the boys happens to get lucky in this lottery, then he may wind up with a payoff that is greater...
than the other's. Someone who comes along at a later date may look at the distribution of toys and conclude that it is unfair, since one boy has ‘more’ than the other. Nevertheless, the Nash solution regards this as fair, because they were each given a fair chance of getting the goods. As a result, a situation that looks unfair from one perspective may seem fair when seen from another. It all depends upon what one takes to be the correct equalisandum.

So before deciding what we mean by equality, we first have to address the more fundamental question, "equality of what?"

8.1 The critique of welfarism

In everyday political discourse, it is conventional to distinguish between "equality of opportunity" and "equality of outcome." This is, unfortunately, not the most fundamental way of dividing up conceptions of equality. Both equality of opportunity and equality of outcome usually assume that the correct equalisandum is welfare, or preference-satisfaction. Equality of opportunity attempts to equalize expected utility, while equality of outcome attempts to equalize payoff. Contemporary philosophers, however, following a highly influential series of articles by Ronald Dworkin, usually draw a distinction between equality of resources and equality of welfare. Resources are understood to be the tangible goods and services that agents use to satisfy their needs, desires and projects – the kinds of things that are typically the object of exchange in a market economy – while welfare is understood to be the satisfaction that agents get, or expect to get, when their needs, desires and preferences are satisfied.

At first glance, welfare egalitarianism appears to have the strongest case. Equalizing just resources seems to be unnecessarily narrow, and limits the egalitarian framework to only the most material products of social cooperation. In order to say that people are truly equal, it is important that their lives all go well, that they each are able to achieve their own goals and projects. Simply handing out material goods in equal-sized bundles seems hardly adequate to ensure this result.

To see how the welfarist intuition can be developed, consider once again the example of the collective farm from Chapter 4. Your economic advisor encouraged you to permit trade, and ultimately to distribute out parcels of land, in order to promote allocative and productive efficiency. But suppose now that, having accepted this solution, you begin to wonder how you can allocate the resources so that the outcome will also be fair. Should you give everyone plots
of land that are exactly the same size, and exactly the same set of gardening tools? What if some people don't intend to use all of their land, or all of those tools? Or maybe some people are better at using one tool or another. Perhaps the distribution of resources should be tailored to people’s specific needs. But if this is the case, how do you even begin to figure out who should get what? After all, as soon as you stop giving out identical bundles to everyone, and start giving out different bundles to different people, how can you decide any more whether the distribution is equal? It would appear that you need some sort of common metric, so that you can measure the value of the bundle that each person receives.

One natural way to approach the problem is to ask whether your distribution, at the end of the day, will make everyone equally happy. The level of satisfaction produced provides, at least in principle, the basis for a common metric for determining the value of a heterogeneous bundle of goods. Furthermore, your ultimate goal is to come up with a distribution that everyone can accept. If the distribution makes everyone equally happy, then it would seem that no one would have reason to reject it. And since the goal of cooperation is to promote human well-being, and secure mutual advantage, it seems reasonable that the advantages conferred by cooperative arrangement should be distributed in such a way as to make everyone equal in this regard.

This is the central intuition underlying welfarism. It is clearly one of the assumptions made by bargaining theory, since the latter defines the distribution problem over a utility space. People's level of well-being is determined by the ability to achieve their own goals and projects, to satisfy their preferences. Since this is what utility represents, this form of welfarism can be seen in the idea that fairness, or equality more generally, should be defined in terms of the utility levels that players are able to achieve under different arrangements. (Although it should be noted that the utility functions used in the previous chapter are cardinal and non-comparable. Thus it is not really the absolute magnitude of satisfaction produced in both individuals being compared, but the size of the increments relative to some other possible increment.)

The central problem with this view is that welfare, defined in terms of preference, cannot be directly redistributed. One cannot take a certain amount of welfare from one person and give it to another. The only things that can be redistributed are material resources, along with certain types of opportunities to achieve welfare (e.g. a job, a spot on the Olympic team, etc). Each individual must then use these resources to produce his or her own welfare. In this respect, the
agent's preferences represent the mechanism through which he or she converts resources into welfare. When an agent is given a bunch of carrots, for example, she can eat them and derive a certain amount of satisfaction from doing so. How much satisfaction she gets is determined by how much she likes carrots (which is what her preference for carrots, and other vegetables, is intended to represent). Society cannot determine her satisfaction level directly, because it has no control over her preferences (assuming, again, that perfectionism is not a live option). Thus society can contribute to her satisfaction only indirectly – by changing the number of carrots available to her.

So even when we set out to equalize welfare, we can only do so indirectly, by distributing resources. This is not in itself a problem for welfarism. The mere fact that we can only distribute resources does not mean that we should only be concerned with equalizing resources. We might choose to distribute resources in such a way that everyone ends up with "same" relative utility level ("relative" reflects the fact that we are not making interpersonal comparisons). Doing so, however, will mean giving out very different bundles of resources to different individuals. It is clear, for example, that giving out equal-sized plots of land to members of the collective farm is not, in general, going to promote equality of welfare. People also have different tastes and temperaments. If one person really likes potatoes, which take twice as much land to cultivate as carrots, the level of welfare obtainable from a certain-sized plot of land will be half that obtainable by a carrot-lover. As a result, the goal of equalizing welfare will quickly come apart from the goal of equalizing resources.

This is where the deeper and more subtle problem with welfarism arises. The rate at which individuals convert resources into welfare will, in general, differ, depending upon the sort of preferences that they have. If we are trying to equalize welfare by distributing out land, it seems that the potato-lover should get more land than the carrot-lover, because satisfaction of the potato-lover's preferences simply requires more resources. The potato-lover has what is called an expensive preference. The odd thing about welfarism is that it suggests that society should shift resources away from people with frugal tastes, and give them to those with expensive ones. Consider the following example, from Dworkin:

Suppose, for example, that a man of some wealth has several children, one of whom is blind, another a playboy with expensive tastes, a third a prospective politician with
expensive ambitions, another a poet with humble needs, another a sculptor who works in
expensive material, and so forth. How shall he draw his will? If he takes equality of
welfare as his goal, then he will take these differences among his children into account,
so that he will not leave them equal shares. (1981, 187).

When there is a **handicap**, such as blindness, the welfarist solution does not seem unduly
problematic (although it is still contestable). However, with expensive tastes it seems clearly
wrong. For example, giving the playboy more than the poet – on the grounds that the poet is
happy with very little – would seem to reward profligacy and penalize frugality.

For a more precise example, consider two individuals, charged with splitting up some set
of goods. Consumption of the good is subject to diminishing returns, so that each additional unit
consumed is worth .9 of what the previous one was worth. Dividing up 40 units of the good
between the two generates the set of Pareto-optimal allocations shown in the blue (diamond)
series in Figure 8.1 below. The Nash solution here has each individual receiving 20 units, for a
utility of about 7.9 each.

However, if the first player happens to become satiated with the good more quickly, then
the value of an additional unit will decrease more quickly for him than it does for player 2. If an
additional unit of the good is only worth .8 of the previous to him, then the Pareto-frontier of
their distribution problem will be given by the pink series in Figure 8.1. The problem is that
while the Pareto-frontier shifts inward, the Nash solution shifts upward, giving player 2 a larger
allocation of goods. In this problem, the Nash solution assigns player 1 only 17 units of the good
(for a utility of 6.1), and player two 23 units (for a utility level of 8.2).
Figure 8.1: Two marginal rates

The perverse consequence is that, in this second distribution problem, player 1 is effectively penalized for being more easily satisfied. When the two players are just as easily pleased (or when they convert resources into welfare at the same rate) they both get the same amount. However, when one player becomes harder to please than the other, the Nash bargaining solution shifts resources toward the one who is harder to please (the Kalai-Smorodinsky solution does the same).

What exactly is wrong with this kind of entitlement? It would be question-begging to say that such a redistribution is unfair, or unjust, since the definition of these terms is precisely what is at issue. However, there are two other obvious objections to this kind of compensatory
arrangement, both of which avoid circularity. The first is an argument from moral hazard, the second points to an objectionable form of unilateralism.

1. Moral hazard. The moral hazard problem is quite straightforward. Acquiring an expensive taste generates a negative externality under regime of welfare egalitarianism. It means that resources must be transferred away from others, and given to the agent who has acquired the expensive taste, in order to maintain equality of welfare. However, the agent is fully compensated for having acquired this new preference. This makes it more likely that people will acquire such preferences, and so herein lies the moral hazard problem.

As Dworkin points out, "expensive tastes (by definition) decrease the total welfare that can be produced from a given stock of resources" (1981, 235). In the same way, a fire damages the stock of buildings available to society. People who are fully insured against fire damage will tend to take fewer precautions to ensure that a fire does not occur, just as people who are fully insured against the costs of satisfying expensive preferences will take fewer precautions to ensure that they don't acquire them. As a result, people are more likely to act in ways that decrease the total welfare of society, simply because they have lost all incentive not to. (In fact, this has the potential to generate not only a suboptimal outcome, but even a race to the bottom. If one person cultivates an expensive taste, and passes all the costs along to others, others may be inclined to develop the same taste, in order to “even things up.”)

One way to think of welfare egalitarianism is as a giant insurance scheme, which protects individuals against almost anything bad that could happen to them – or more specifically, one that insures them against unhappiness. As a result, there is the danger that people will lose whatever incentive they have to ensure that they get an adequate level of welfare out of their resources. They might lose the incentive that people normally have to not develop expensive tastes.

As Dworkin correctly notes, the moral hazard objection is really an efficiency problem, not a problem with the theory of justice per se (1981, 235). If everyone exercises diminished caution in the way that they use resources, then everyone will still be equal, it is just that they will be equal at a lower level of welfare than would be obtainable without the equalization scheme. Since equality and efficiency are at least in principle distinct, this is not really an objection to any particular conception of equality. Nevertheless, it suggests that any “welfarist”
conception of equality is going to interact with the principle of efficiency (when combined into a general theory of justice) in highly undesirable ways.

2. **Unilateralism.** The other objection to the cultivation of expensive tastes is that it allows individuals to make private choices that have important consequences for the welfare of others, yet prevents these others from having any say in the matter. Thus it allows one individual, as Arthur Ripstein puts it, to “unilaterally set the terms of interaction” in an objectionable way. By externalizing the cost of expensive preferences, agents are making decisions that impose costs upon others, without having to take these costs into consideration in their deliberations, and without having to consult those affected by their choice. Dworkin writes:

Imagine that a particular society has managed to achieve equality of welfare in some chosen conception of that ideal. Suppose also that it has achieved this through a distribution that in fact (perhaps just by coincidence) gives everyone equal wealth. Now suppose that someone (Louis) sets out deliberately to cultivate some expensive taste or ambition he does not now have, but which will be expensive in the sense that once it has been cultivated he will not have as much welfare on the chosen conception as he had before unless he acquires more wealth... Louis has a choice. He may choose to keep the presently equal resources I said he had, and settle for a life with the enjoyment he now has but without the tastes or ambitions he proposes to cultivate. Or he may keep his present resources and settle for a life that he deems more successful overall than his present life, but one that contains less enjoyment. [But] it is quite unfair that he should have a third choice, that he should be able, at the expense of others, to lead a life that is more expensive than theirs at no sacrifice of enjoyment to himself just because he would, quite naturally, consider that life a more successful life overall than either of the other two (1981, 237).

Thus Louis is entitled to make a choice over the first two options, but is not entitled to the third. "Louis knows, or at least ought to know, that if he cultivates some expensive taste in a society dedicated to equality of enjoyment, for example, and is compensated, then that will decrease the enjoyment available for others" (1981, 237). Dworkin's does not claim that Louis
should not have the entitlement to acquire the expensive preference. Acquiring the preference is only objectionable if it involves passing the cost along to others. Making a choice that decreases the level of enjoyment available to others allows him to unilaterally set the terms of interaction. Thus he should be entitled to the choice, but he should also be required to bear the costs.

Thus the problem with welfarism is that it violates the idea that cooperation should be governed by terms of interaction that all could accept (or at very least, the idea that individuals should not be able to unilaterally dictate these terms). Welfare cannot generate entitlement, because individuals are able to modify their preferences, and thus control the level of welfare they get out of any given bundle of resources. If they can modify their preferences, and these preferences in turn generate entitlements, it gives individuals the ability to unilaterally expand their entitlements. This allows them to modify the terms of cooperative interaction \textit{ex post}. And certainly any individual could reasonably reject a cooperative arrangement that allowed any one person to unilaterally modify the distributive consequences of that arrangement at a later date.

These arguments suggest that the only acceptable principle of equality will be one that favours cooperative arrangements that are immune to subsequent unilateral revisions. According to Dworkin, this means that the only general principle for deciding distributive questions will be one that equalizes the distribution of \textit{resources}, because only resources provide a basis for "tamper-proof" cooperative agreement. Under such a scheme, resources will be distributed equally, and individuals will then "convert" them into welfare on their own. The level of welfare they achieve will be determined by the type of preferences they have, the level of prudence that they exercise, and so on. People with expensive tastes will get less satisfaction, in general, than those with inexpensive ones. Thus resource egalitarianism has the effect of "internalizing" the costs associated with the acquisition of expensive preferences, thereby discouraging individuals from acquiring them, and preventing individuals from imposing the costs of these preferences on others.

\textbf{8.2 Superfairness}

If we accept the idea that the correct equalisandum for the conception of fairness is resources, not welfare, then how is the principle of equality to be specified? The immediate problem is that resources, taken by themselves, are incommensurable. Apples cannot be
compared to oranges. One of the attractions of utility is that it provided, at least in principle, a common metric for comparison. Furthermore, one of the key motivations for defining welfare in terms of utility, or subjective preferences, is that it permitted the formulation of a principle of equality that would be neutral with respect to people’s different conceptions of the good. In a pluralistic society, we cannot expect to achieve consensus about the value of different goods and services, because different people with different life projects will have different uses for them. For example, the “value” of a car cannot be determined objectively, because some people want cars, some people don’t, and many others would be happier living in a world in which there were no cars. Thus if one person, under a particular social arrangement, gets a car, while someone else doesn’t, it is difficult to say whether this makes them unequal – it all depends upon how much either one of them wants a car. One of the attractions of welfarism is that it promises a simple solution to this problem: just look at how much happier the arrangement makes the two people. If one person doesn’t want a car, and someone else does, then the fact that this second person receives one and the first does not many not violate equality, since they both may wind up equally happy at the end of the day. The resource egalitarian, on the other hand, cannot appeal to “happiness” or satisfaction as a metric, and so must find some other way of determining whether allocations are equal. Luckily, there are options.

Consider a very mundane context, and a relatively simple distribution problem. Suppose you are the executor of an estate, and you must find some way to divide the deceased’s possessions that will give each heir an equal share. It is impossible to achieve this by mechanically dividing of the property into identical bundle (e.g. there is only one clock, you do not want to divide up a dinner service, etc.), and so you cannot “equalize” simply by giving everyone exactly the same thing. It may occur to you to divide up the items by attaching a monetary value to them, then giving everyone a bundle worth the same dollar amount. But this is also unsatisfactory, since it incorporates inessential factors into the deliberations (viz. how much people in society at large want the items), and excludes essential ones (viz. sentimental value to the heirs). If one heir has a great sentimental attachment to something that also has a high monetary value, but that the others don’t want, a distribution that equalizes cash value would be manifestly unfair to this individual. He would have to pay much more for the item than was warranted by the actual cost imposed upon the other beneficiaries.
One way of resolving the problem is to use the following procedure. Suppose you ask everyone questions about what they want, then divide up the goods into a set of bundles. If there are three beneficiaries, you will make three bundles, etc. Each beneficiary then writes down on a piece of paper which bundle he or she wants, and submits this request to you. If the set of requests is \textit{compossible} (i.e. each beneficiary requests a different bundle, and so all the requests can be satisfied simultaneously), then the bundles are distributed out, each one given to the person who requested it. However, if any two people request the same bundle, then none of the goods are distributed. You then consult the preference schedules again, reorganize the bundles, and the entire procedure is repeated. No distribution is made until an allocation is achieved that generates a mutually consistent set of requests.

What you are doing here is constructing what economists and philosophers call an “envy-free” distribution of the estate. The attractions of this type of allocation are easy to understand. In bargaining theory, the central egalitarian principle is the Symmetry axiom, which expresses the intuition that a cooperative arrangement would be acceptable to all only if no one wanted to switch positions with anyone else. With resources, it is possible to apply the same idea in a fairly direct way. We can say that a proposed distribution of resources is fair if no one prefers anyone else’s share to her own (so that if each had to request one bundle, each would choose the one that he or she receives). If we think of “positions” as particular bundles of resources, then individuals will not want to change positions with each other whenever they prefer their own position to that of anyone else’s.

This is the central idea underlying the principle of \textbf{envy-free allocation}, although the term “envy” is misleading. The concept of envy here is not that of an emotion that makes you unhappy, or a type of interdependent preference structure (as in §1.5) but a merely a particular pattern of preference over outcomes.

\textbf{Envy}: A distribution of \textit{n} commodities is said to involve envy by player 2 of the share obtained by player 1 if player 2 would rather have the bundle of commodities received by player 1 under this distribution than the bundle the distribution assigns to player 2 (Baumol 1986, 19).
This is actually a very old idea. The basic principle is stated with admirable clarity by Hobbes in the *Leviathan*. Hobbes wanted to show that people in the state of nature would be roughly equal to one another in both physical and intellectual endowment. Equality of physical strength was fairly easy to establish, because it is easy to make interpersonal comparisons. Intellectual ability, on the other hand, presented something of a measurement problem (in the dark ages before standardized testing). Hobbes argued, therefore, that the most reliable indication that people are roughly equal in their intellectual abilities is the tendency everyone has to think that they’re smarter than everyone else. As Hobbes put it, “there is not ordinarily a greater signe of the equall distribution of any thing, than that every man is contented with his share” (1651, 87).

The mechanical division of an estate into bundles that each heir selects is one way of constructing an envy-free allocation. Once the distribution of the bundles is complete, no one will envy anyone else's share. But the absence of envy is not what makes the procedure attractive, from the standpoint of justice. The important point is that an envy-free allocation can be unanimously endorsed, because everyone likes the bundle that he or she gets under the proposal arrangement better than anyone else's. If two people want the same bundle of goods, then they may start to fight over it. If no two people want the same bundle, then there is nothing to fight over. No one has any incentive to reject the arrangement, because they all prefer the bundle they get to the bundle anyone else gets. And no one can complain later that he or she was treated unfairly by the distribution, Thus the envy-freeness principle is consistent with idea that cooperative arrangements must be agreed to by all.

The procedure outlined above is simply one particular mechanism for achieving an envy-free allocation, and a slightly awkward one at that. A robust mechanism could be created by setting up an auction among the heirs. Each could be given a equal share of some special "currency" created by the executor. The executor can then auction off the resources. When a set of market-clearing prices emerge, these will reflect the intensity with which the various heirs want the goods. The bundles of goods that each heir eventually purchases will be worth exactly the same amount (when expressed in terms of the special currency that the executor has created). As a result, the allocation will be envy-free, because each bundle is worth the same amount. (If one heir preferred someone else's bundle to her own, then she would have purchased that bundle – not her own.)
The envy-freeness concept can be used to define the following principle, which is (confusingly) referred to as fairness:

**Fairness:** A distribution is fair if and only if no individual prefers the share that is received by any other individual to his or her own.

The most important thing to note about this fairness principle is that, like the principle of symmetry in bargaining theory, it does not require interpersonal comparisons of utility. The notion of preference is entirely subjective. Thus an allocation of goods that is fair for one group of people would not remain fair if the members of the group changed, because the new group would have different preferences. Again, this can be seen clearly in the case of estate divisions. While the fact that one person gets grandmother's clock may not inspire envy in the other heirs, because they don't like the sound of its chimes, a different set of heirs might feel quite differently about it. Furthermore, a distribution to is perfectly fair, from the standpoint of the heirs, might seem outrageously unfair to an uninformed observer (for instance, the monetary value of the bundles might differ considerably).

From now on, the term fairness will be used to refer to this specific conception of fairness-as-envy-freeness. Following Baumol (1986), we can also define a slightly stronger principle:

**Superfairness:** A distribution is superfair if and only if each participant receives a bundle that is strictly preferred by that individual to the bundle received by anyone else.

This says that an allocation must not only generate no envy, but that each individual must positively prefer his or her share to anyone else's. Under the fairness principle, an individual could be indifferent between her share and someone else's, and so would still be willing to accept certain rearrangements of the bundles. Under superfairness, the agent has a strict preference for her own bundle, and so would reject any arrangement under which she received anyone else's. This slight difference can be significant. For example, one way of guaranteeing an envy-free allocation is, when possible, to create identical bundles. Here there will be no envy, because
every bundle is identical (inducing indifference in all agents). Call this the equal allocation. It will always be fair, but never superfair.

When there is only one type of resource to be distributed, the only envy-free allocation is an equal one (and there are no superfair distributions). So in the distribution problem shown in Figure 8.1, the only fair allocation is to give each player 20 units (regardless of how much either of them likes the good in question). Similarly, if people have identical preferences, then only fair allocations will be available, not superfair ones. With identical preferences, the contents of the bundles may differ, because individuals may be indifferent between different mixes of goods. But if one of the bundles is such that one person prefers it over all the others, then everyone will prefer it over all the others, and so the resulting allocation could not be envy-free. Thus no allocation will be superfair.

Thus the fairness principle prescribes strict equality of resources in most simple allocation cases. This I take to be in accord with our general intuitions, and our experiences dividing up candy among children and such things. The more robust character of the principle becomes evident when it is applied to more complex cases, involving different individuals with different preferences, and a mixed bundle of commodities. Here the principle can be used to resolve problems that could not be handled using any mechanical method of dividing the goods. But there are still some problems.

8.3 Indeterminacy problems

It is also important to notice that neither fairness nor superfairness offers as determinate a solution to distribution problems as the bargaining solutions presented in the previous chapter. For instance, Nash's solution to the toy-distribution problem is superfair. Unfortunately, there are many other superfair allocations. Table 8.1 shows the two bundles that Nash's solution divides the toys into, along with their utility for each boy. This solution is superfair, since each boy prefers the bundle that he receives to the one that the other gets. Bill prefers having the "box, pen, toy and knife" bundle to the "book, whip, ball, bat, hat" bundle, and Jack has the opposite preference.
Table 8.1

In Bill's case, the very high utility associated with his bundle is largely due to his very strong desire for the pen. This creates something of a problem, since it means that the toy or the box could be removed from his bundle, without generating any envy. Consider the sets shown in Table 8.2:

<table>
<thead>
<tr>
<th>Bill’s utility</th>
<th>[box, pen, toy, knife]</th>
<th>Jack’s utility</th>
<th>[book, whip, ball, bat, hat]</th>
</tr>
</thead>
<tbody>
<tr>
<td>24</td>
<td></td>
<td>5</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 8.2

This allocation is efficient and superfair. Similarly, Jack could give Bill the hat, as shown in Table 8.3:

<table>
<thead>
<tr>
<th>Bill’s utility</th>
<th>[pen, toy, knife]</th>
<th>Jack’s utility</th>
<th>[box, book, whip, ball, bat, hat]</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td></td>
<td>4</td>
<td>14</td>
</tr>
</tbody>
</table>

Table 8.3

Thus superfairness does not pick out a single solution, but rather a set of possible distributions. Figure 8.2 shows the set of superfair allocations, as a subset of the utility space of the two boys:
This is a problem, because it introduces a significant element of arbitrariness into the solution. Imagine, for example, the division of an estate between two heirs, both of whom have one set of items that they are very strongly attached to – so strongly that they will automatically prefer any bundle which contains it. If $X$ is the set of items that player 1 wants, $Y$ is the set of items that player 2 wants, and $Z$ is the rest of the estate, then $\{X \cup Z, Y\}$ and $\{X, Z \cup Y\}$ are both superfair allocations. This suggests that it is of no consequence, from a distributive standpoint, how the bulk of the estate ($Z$) is to be allocated, and that solutions in which it is all given to one person are perfectly fair. Not only does this seem arbitrary, but it also contradicts our sense that $Z$ should be split evenly between the two heirs.

A variety of proposals have been made to resolve this problem. One is to treat the region of superfair allocations as another distribution problem, and then pick out the superfair solution to that problem. The "ultimate" solution might then be the limit of such a series of recursive applications of the superfairness principle. Another possibility is simply to use a mechanical formula to divide up the resources in the superfairness region. In any case, in order to explore
either of these two options, it is helpful to have a graphic representation of how the superfairness principle operates.

### 8.4 A graphical representation

The key idea underlying the superfairness solution concept is that if there is no envy, then no player has an incentive to switch places with anyone else. This no-place-switching principle can be illustrated with an Edgeworth box diagram, the construction of which is illustrated in Figures 8.3 through 8.8.

![Edgeworth box diagram](image.png)

**Figure 8.3 An indifference curve**

Figure 8.3 shows an **indifference curve** for an agent, in this case Bill. Indifference curves were introduced in Chapter 4 in order to show how much of one particular good an agent would be willing to give up in order to receive some quantity of another goods (both supply and demand curves are indifference curves). This kind of graph can be used to represent preferences, since any point to the NE of any point on the indifference curve is preferred by the agent to every point on that curve. Thus it is said that any point on a "higher" indifference curve is preferred to any point on a "lower" one.

Suppose now that instead of being interested in the tradeoffs that an individual would be willing to accept between two goods, we are interested in knowing how the individual would respond to proposals for dividing up some fixed quantity of these goods between himself and
another individual. We want to know which divisions will lead Bill to want to switch places with
the other person. There is a very simple procedure for determining this.

For simplicity, we can begin by rescaling the problem so that 1 unit of carrots and 1 unit
of potatoes are to be divided up. Under every proposal to allocate the goods, one bundle will
contain \( x \) carrots and \( y \) potatoes, the other will contain \((1-x)\) and \((1-y)\). For any proposed division,
the first thing to do will be to plot out both bundles, \((x,y)\) and \((1-x,1-y)\) (the second point can be
referred to as the "image" of the first, because within the space delimited by \((0,0)\) and \((1,1)\), it is
the exact mirror image). Figure 8.4 shows three proposed allocations. The first allocation gives
Bill \( p \), and hence the other person \( p' \), the second gives him \( q \), the other person \( q' \), while the third
gives him \( r \), and the other person \( r' \).

![Figure 8.4 Detecting envy](image_url)

In order to determine whether these proposed allocations make Bill envious, one begins
by drawing the indifference curve that passes through the bundle that Bill is to receive. (For
convenience, \( p, q \) and \( r \) are all assumed to be on the same curve in Figure 8.4). There are then
three possibilities:

- The image lies above the curve: Bill prefers the second bundle to the first, and so the
  proposed distribution induces envy. This is illustrated by \( q \) and \( q' \) in Figure 8.4. The fact
  that \( q' \) lies on a higher indifference curve for Bill than \( q \) reflects the fact that Bill would
  prefer to receive a bundle that contains \( 1-x \) carrots and \( 1-y \) potatoes, than one that
contains $x$ carrots and $y$ potatoes ($x$ and $y$ being, in this case, the quantities of those goods in bundle $q$).

- The image lies below the curve ($r$ and $r'$): Bill prefers the first bundle to the second, and so the allocation does not make Bill envious.
- The image lies on the curve ($p$ and $p'$): Bill is indifferent between the two allocations.

Rather than plot out each allocation individually, the set of "images" can actually be drawn as a set of "shadow" indifference curves, each of which precisely mirrors the agent's actual preferences. Thus curve $b_2$, shown in Figure 8.5, represents the allocations that the other person gets, for each of the points on $b_1$.

![Figure 8.5 Fairness threshold points](image)

The points at which the indifference curve intersects with its image represent the set of allocations that make Bill indifferent between what he and the other receive. Any allocation that lies on the indifference curve *between* these two points will be one that leads Bill to envy the other player's bundle. Anything on the line beyond these two points *may* be fair, because it does not create any envy in Bill. Thus these two points determine Bill's **fairness threshold** for points on this particular indifference curve $b_1$. Lying on the indifference curve *beyond* this threshold (i.e. further from the origin) is a necessary but not sufficient condition for an allocation to be fair.
(It is necessary because any point on $b_1$ between these two points represents an allocation that will make Bill envious. It is not sufficient because nothing has been said yet about how the other player feels, and a fair allocation must not create envy in either of the players.)

These two points only specify the fairness threshold for the points lying along the same indifference curve ($b_1$). But we can go on to specify where the fairness threshold lies for all possible allocations, simply by picking out the fairness threshold for all of the agent's indifference curves. This generates a line, as shown in Figure 8.6. This can be referred to as the fairness frontier, since any point on it is potentially fair (i.e. does not inspire envy in Bill), and any point above it (i.e. to the NE of it) is potentially superfair.

![Figure 8.6 Bill's fairness threshold](image)

**Figure 8.6 Bill's fairness threshold**

Again, it is important to keep in mind that points above Bill's fairness frontier are only potentially fair, since they represent only the set of allocation that do not make Bill envious. In order to determine what is actually fair, one must begin by constructing a similar diagram for the second person (which will be different, in cases where that individual’s preferences – and hence indifference curves – are different). Figure 8.7 shows a possible fairness frontier for Bill's opponent, Jack.
Figure 8.7 Jack's fairness threshold

In order to determine the set of superfair allocations, one can then simply superimpose the "image" of one person's fairness frontier onto the other's. This can be done by flipping Jack’s diagram along the diagonal axis, so that the upper-right hand corner of Bill’s allocation diagram is the origin of Jack’s (i.e. if Bill gets the full 1 unit of carrots, that is equivalent to Jack getting zero, etc.) This is known as an Edgeworth box diagram.

When the image of Jack's fairness threshold is superimposed on Bill's, the area where the two sets of "potentially fair" allocations intersect represents the set of actually superfair allocations. In this region, each player positively prefers his own bundle to the one that the other gets.
8.8 Superfair allocations between Bill and Jack

The point where the two fairness thresholds (i.e. the lines) intersect represents the allocation that is exactly fair – neither player envies the other's bundle, but neither prefers his own. This will usually be the equal allocation. The intersection of the two spaces beyond the fairness threshold represents the set of superfair allocations. This graphical representation makes it clear why the superfairness standard can be indeterminate (i.e. can pick out more than one allocation). In this case, Bill like potatoes somewhat more than Jack does (relative to carrots), and so both individuals are willing to tolerant a certain amount of deviation from strict numerical equality, in the case of distributions that give Bill more potatoes than carrots.

Figure 8.8 also shows why, when agents have exactly the same preferences, there will be only fair solutions. When preferences are identical, the image of one player's fairness frontier will be the same as the other's actual frontier, and so the two lines will lie on top of each other when superimposed. As a result, there will be a set of fair solutions (each point on the line), and no superfair ones.

This completes the task of specifying an appropriate fairness principle for the division of resources. Even though the principle does not pick out a unique solution, it does specify a set of distributions that should be acceptable to all – in the sense that no one will have any desire to
switch places with anyone else. (In the same way, the Pareto-efficiency principle does not pick out a unique solution.) This may not be the last word, but it does represent a significant constraint. It takes us a long way toward the goal of specifying principles that can be used to secure agreement on the division of a cooperative surplus.

8.5 The question of scope

The discussion so far has proceeded on the assumption, stated in the introduction of the text, that the primary function of social institutions is to secure certain cooperative benefits, and that the purpose of a theory of justice is to provide principles for determining the magnitude and distribution of these benefits. Implicit in this construction is the further assumption that the scope of these principles extends only to these benefits. Thus with respect to equality, it assumes that what needs to be equalized are the benefits of cooperation; it says nothing about what should be done with respect to other forms of inequality that arise from various aspects of the human condition.

Thus Rawls, for example, argues that the proper equalisandum for a theory of justice is a class of resources that he refers to as primary goods. These are goods that serve as all-purpose means to the realization of one's projects, whatever those projects may be. However, within the category of primary goods, Rawls distinguishes between primary social goods and primary natural goods. The former are goods typically produced through cooperation, e.g. wealth, education, etc., while the latter are not produced through cooperation, but belong to the “natural” endowment of the agent, e.g. health, motivation, strength, skill. The principles of justice, in Rawls's view, call for an egalitarian distribution of the former, but not the latter. In other words, it is important that social institutions be immunized against the effects of natural inequality, so that the effects of the unequal distribution of natural primary goods do not wind up being amplified by an unequal distribution of social primary goods. However, there is no obligation on the part of society, grounding in principles of justice, to redress natural inequality, or to compensate individuals who receive a below-average endowment of natural primary goods.

From this perspective, which we can refer to as narrow-scope egalitarianism, it is easy to see why welfarism encounters difficulties. The fundamental problem is that welfare is not one of the products of cooperation. It is, rather, produced by individuals, through their preferences, using various systems of social cooperation to provide the necessary “inputs.” Thus cooperation
does an enormous amount to promote *improvements* in the welfare of all, but it does not directly produce welfare. This is why, within welfarist forms of egalitarianism, individuals wind up having an objectionable form of control over the system of cooperation – in effect, the ability to manipulate their entitlements by cultivating expensive preferences and externalizing the cost. Resource egalitarianism, by contrasts, confines the redistributive system to the actual fruits of cooperation – the goods and services that individuals produce. (Of course, the terms “resources” needs to be understood here in a very broad sense – not just the material goods that are exchanged in the market economy, but also public goods, club goods, and other goods that are currently produced only as externalities.)

There is, however, another important current of egalitarian thinking that rejects the narrow-scope egalitarian's confinement of the principles of justice to the fruits of cooperation. Dworkin, for instance, believes that the principle of equality demands not merely that we immunize social institutions from the effects of natural inequality, but that the pattern of distribution be such as to (at least partially) redress this natural inequality. Thus he brings the distribution of natural “resources” or talents within the *scope* of distributive justice. We may refer to this as **wide-scope egalitarianism**. Of course, the elements of an individual's natural endowment tend not to be fungible, and so cannot be redistributed. Thus he calls for a system of differentiated transfers of the social product, in amounts inversely related to individual's natural endowment (e.g. “greater resources must be spent on the education of the less rather than the more intelligent.” [Rawls 1971, 101]). When this is done, each individual's total allocation of natural and social resources will ideally come out to be the same (the full story is slightly more complicated, but this is the general idea). Figure 8.9 illustrates what such a distribution would look like.
The positive case for wide-scope egalitarianism is usually made through appeal to some version of what Rawls call “the principle of redress,” according to which “undeserved inequalities call for redress; and since inequalities of birth and natural endowment are undeserved, these inequalities are somehow to be compensated for” (Rawls 1971, 100). The motivation for the principle of redress usually lies in a set of moral intuitions concerning the role that luck and responsibility should play in determining people’s entitlements. The typical sort of argument runs like this: A person who is born with a handicap faces dramatically reduced chances for attaining any of the elements that we consider important for a successful or even enjoyable life. Yet this person has done nothing to deserve this fate. Because he does not deserve this condition, he cannot be held responsible for its effects – be they poverty, unhappiness, incapacity, or what have you. Similarly, people who get lucky and receive an exceptional natural endowment have done nothing to deserve the associated rewards. Thus society is under an obligation to effect a transfer from those who make undeserved gains to those who suffer undeserved losses. Such a transfer is justified even if it disadvantages some people to the point where they become net losers from their participation in society. After all, these people have
done nothing to deserve their superior initial endowments, and so have no moral grounds to resist the appropriation.

The broader ideal that informs this view is one of a society in which reward is exactly proportional to desert. Under such conditions, the distribution of the social product would be justified in a very strong sense of the term. If anyone asks why so-and-so has such-and-such, it would be possible to tell a story that would justify that precise endowment. Another way of conceiving of this is to imagine a society that is completely immunized against luck (Cohen, 1989, 931). Inequality would be permissible when they arose as a result of the choices people had made, but any inequalities due merely to their circumstances (or “brute luck”) would be eliminated by the social transfer.

There is a large, extremely complex philosophical literature on this argument for egalitarianism (which has come to be known as “luck egalitarianism”). The moral intuitions described above are, in general, not decisive. There is some danger of equivocation in the suggestion that because the individual has done nothing to deserve his or her natural endowment, that “society” as a whole should assume the burden (by providing full compensation). After all, society is just a shorthand way of referring to “other people.” From the fact that one individual is not responsible it does not follow that someone else must be. Such an inference ignores the possibility that, in many cases, no one is responsible. The question is therefore what we should do with losses which are not the product of anyone’s choice. Should they lie where they fall, or should they be collectivized? The case for collectivization of such losses must be made on its own terms – it does not follow directly from the mere fact that the individual most directly affected is not responsible.

It should also be noted that the extension of principles of equality to encompass natural endowments leads quite quickly to a series of problems, which critics usually identify with egalitarianism, but are actually a product of the scope with which egalitarian principles are being applied:

1. Existence problems. If egalitarian principles apply only to the social product, then we can rest assured that there will always be a feasible egalitarian allocation. But as soon as we include natural endowments in the distribution problem, this assurance disappears. Because we cannot
change the natural distribution, there will often be nothing that we can do to produce an allocation that satisfies the principle of equality. The distribution problem will have no solution.

For example, suppose we attempt to correct natural inequality by establishing an envy-free distribution of both natural and social primary goods combined. There is no reason to think that this will be possible. Many people are prepared to make very extreme sacrifices in order to overcome natural handicaps. Consider the blind person who would “give anything” to have his sight back. Taken literally, this means that no transfer of primary social goods, no matter how large, would be enough to make this person’s total bundle of primary goods (which will now include the absence of vision) as desirable as that of a sighted person (which will include the natural good of vision). If we stop for a moment to consider how many “inconsolable” victims of bad luck there are, we may want to think twice about imposing upon society the obligation to channel arbitrarily large quantities of the social product toward such individuals. (The problem is even more severe for welfarists than for resourcists, since they must have something to offer those who suffer all forms of human tragedy, such as the murder of a loved one.)

2. Money pits. Even if the presence of severely under-endowed individuals does not make an egalitarian allocation impossible, it can still create an obligation to transfer a staggering portion of the social product to such persons. Certain individuals may become the human equivalent of money pits, consuming an enormous quantity of resources in return for very little improvement in overall condition. This is often thought to be an unattractive conclusion, given that it can “have too depressive an effect on the welfare of everyone else in society” (Cohen 1989, 911). However, this concern is often brusquely dismissed on the grounds that, as G.A. Cohen puts it “it hardly represents an egalitarian objection” (1989, 991). In particular, it is often thought to be merely an efficiency concern, and thus not a problem for the principle of equality, strictly construed. However, it is far from obvious that there are no issues of equality at stake in these examples. It is true that the central problem is one of waste (and thus of efficiency). For instance, vast quantities of resources can be expended to extend the life of the sickest person in the hospital by one day, instead of, for example, dramatically reducing the waiting list for procedures that merely improve the quality of life of others. But at the end of the day, the sickest person still dies. Thus it seems like a waste to have channeled so many resources into this person’s care. It would require a much more elaborate argument than Cohen provides to show
that these intuitions reflect entirely non-egalitarian considerations. While our efficiency sensibilities are offended by the waste of resources, our egalitarian sensibilities are also injured by the sight of one individual “hogging” too large a share. It is often felt that egalitarian principles should not permit individuals to impose unreasonable burdens upon others. Does not the requirement that one individual receive the lion’s share of the social product represent precisely such an imposition? Does it really treat people with equal concern to entirely sacrifice the life projects of one for the sake of another?

3. Slavery problems. Wide-scope egalitarianism also has the potential to produce very disturbing interferences with individual liberty, up to and including the so-called “slavery of the talented.” (Dworkin, 1981). This problem is in fact just the flip side of money pit problem. Some people may have such a generous natural endowment that they have essentially no entitlement to any of the social product. Thus they may have no material incentive to work, since if they do work, everything they produce will have to go to others, in order to prevent an exacerbation of the initial inequality.

This raises obvious issues about incentives. But it creates far more serious problems when one realizes that any plausible conception of equality is going to have to include leisure as one of the goods that makes up an individual’s allocation. Thus if we leave individuals with a generous natural endowment free to choose their own hours of work, many will choose not to work. But this is equivalent to increasing the quantity of leisure that they enjoy (they are choosing, in effect, to spend their time producing a non-fungible good, in order to avoid the social transfer), which also represents as exacerbation of social inequality. Thus wide-scope egalitarianism carries with it the *prima facie* implication that these individuals are not entitled to make this choice.

This is an unobserved consequence of many wide-scope egalitarian schemes, often ones that seem quite innocuous. For example, lump sum taxes are often proposed as a way of implementing an egalitarian distribution. Individuals with greater natural talent will naturally face a higher tax bill than others. However, these taxes cannot be imposed in the form of an income tax, because of the leisure problem. Thus they will have to take the form of head taxes. This means that highly talented individuals will be forced to work in order to pay their tax bill. Furthermore, the bill may be so large that they are forced to work in their most productive
employment (Van Parijs 1995, 64). In this case, Robert Nozick’s (1973) characterization of taxation as “forced labor” seems much less far-fetched.

4. Leveling down. It tends to be assumed in discussion of egalitarianism that the equal allocation is to be achieved through measures that “lift up” those at the bottom. However, there is no reason in principle why it could not also be achieved by “leveling down” those at the top, without providing any benefits to those at the bottom (Parfit, 1997). This is an especially serious problem for wide-scope egalitarianism, given their preoccupation with non-fungible natural endowments. Even though natural abilities are not transferable, it is possible to handicap people in various ways, so that they are not able to benefit from or enjoy their endowment. If it is not possible to achieve equality among individuals through transfers, why not do so by eliminating some of the underlying natural endowments? We could, as Kurt Vonnegut suggests (in “Harrison Bergeron” (1998), his vicious but not-entirely-mistaken parody of egalitarianism) force the beautiful to wear masks, saddle the graceful with weights, and systematically irritate the intelligent in order to prevent them from concentrating. Because this represents a pure loss to those at the top, with no compensating benefits for those at the bottom, there are obvious efficiency arguments against such arrangements. But wide-scope egalitarianism appears to suggest that such transformations enjoy prima facie normative superiority.

These four problems can be summed up with a very simple graphical illustration (Figure 8.10). In contrast to Figure 8.9, this one shows a case in which there is simply not enough social product available to make up for the difference in natural endowment. Depending upon how one looks at it, one might say that individual 4 makes equality impossible, because of her exceptional natural endowment, or that individual 3 makes it impossible, by consuming so much of the social product in order to make up for her exceptionally poor natural endowment. Either way, it is the magnitude of the natural inequality that makes it impossible for society to achieve an egalitarian allocation.
If one adopts the allocation of the social product shown in Figure 8.10, on the grounds that it is “as close as possible” to equality, the money pit and slavery problems are an immediate consequence. Furthermore, since individual 4 is clearly the “tall poppy,” the comprehensive egalitarian would appear to be committed *prima facie* to “leveling down” – eliminating the portion of that individual’s natural endowment that exceeds the total allocation enjoyed by the others (although this would probably be trumped by efficiency considerations in a “complete” theory of justice).

In order to avoid such problems, the discussion that follows will be based upon the assumption that the principle of equality will be applied in a narrow-scope fashion, to the benefits of cooperation only. Thus no one will ever “lose out” in some absolute sense from their participation in institutions that are organized in accordance with egalitarian principles. They may not do as well as they might have under some other possible institutional arrangement, but because the principle of equality is used only to divide up the cooperative surplus, everyone is

---

**Figure 8.10 Existence problem**

Social product

<table>
<thead>
<tr>
<th>Individual</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social product</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural endowment</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

If one adopts the allocation of the social product shown in Figure 8.10, on the grounds that it is “as close as possible” to equality, the money pit and slavery problems are an immediate consequence. Furthermore, since individual 4 is clearly the “tall poppy,” the comprehensive egalitarian would appear to be committed *prima facie* to “leveling down” – eliminating the portion of that individual’s natural endowment that exceeds the total allocation enjoyed by the others (although this would probably be trumped by efficiency considerations in a “complete” theory of justice).

In order to avoid such problems, the discussion that follows will be based upon the assumption that the principle of equality will be applied in a narrow-scope fashion, to the benefits of cooperation only. Thus no one will ever “lose out” in some absolute sense from their participation in institutions that are organized in accordance with egalitarian principles. They may not do as well as they might have under some other possible institutional arrangement, but because the principle of equality is used only to divide up the cooperative surplus, everyone is
guaranteed to do better by participating in the institution than he or she could do by going it alone.

**Key words**

brute luck
envy
envy-free allocation
equal allocation
expensive preferences
fairness
fairness threshold
fairness frontier
handicaps
luck egalitarianism
option luck
primary goods
resource egalitarianism
superfairness
welfare egalitarianism