

Dworkin's auction

Joseph Heath

University of Toronto, Canada

abstract

Ronald Dworkin's argument for resource egalitarianism has as its centerpiece a thought experiment involving a group of shipwreck survivors washed ashore on an uninhabited island, who decide to divide up all of the resources on the island equally using a competitive auction. Unfortunately, Dworkin misunderstands how the auction mechanism works, and so misinterprets its significance for egalitarian political philosophy. First, he makes it seem as though there is a conceptual connection between the 'envy-freeness' standard and the auction, when in fact there is none. Second, he fails to appreciate how idealized the conditions are that must be satisfied in order for his results to obtain. This leads him to draw practical conclusions from the thought experiment that do not follow, such as his claim that the principle of equality generates a presumption in favor of the market as a mechanism for the distribution of resources. The result is that Dworkin saddles resource egalitarianism with a set of commitments that are, in fact, inessential to that view.

keywords

Ronald Dworkin, envy-freeness, superfairness, efficiency, resource egalitarianism

Ronald Dworkin's argument for resource egalitarianism in his (*Sovereign Virtue* henceforth referred to as SV) has as its centerpiece a now-famous thought experiment, involving a group of shipwreck survivors washed ashore on an otherwise uninhabited island.¹ They decide to divide up all of the resources on the island equally. The principle of equality that they seek to satisfy is the 'envy-freeness' standard: the division is equal if no one prefers the bundle received by someone else to his or her own. In order to achieve such a division, they devise a competitive auction for all the resources, and then give each survivor an equal number of clamshells to use as currency. After the auction is run, the entire island

gets divided up in accordance with the bids, and each survivor is happy with his or her portion. Harmony reigns.

This thought experiment is one of the most widely misunderstood pieces of reasoning in contemporary political philosophy. Most significantly, Dworkin himself misunderstands how the auction mechanism works, and misinterprets its significance for both the practical and theoretical aspects of the theory of justice. First, there is confusion about what the auction accomplishes in the broader resource-egalitarian scheme. Dworkin makes it seem as though there is some conceptual connection between the envy-freeness standard and the auction, when in fact there is none. The auction simply takes a prior envy-free allocation and makes it Pareto efficient. Second, Dworkin fails to appreciate how strict the conditions are that must be satisfied in order for the results that he derives to obtain. Thus, he draws practical political conclusions from this thought experiment that simply do not follow. In particular, he claims that the principle of equality generates a presumption in favor of the market as a mechanism for the distribution of resources, when in fact his argument, correctly understood, implies no such thing.

These issues of interpretation would not be so serious, were they simply localized problems in Dworkin's particular version of resource egalitarianism. Unfortunately, misunderstandings of the auction mechanism have spread beyond Dworkin's work, so that they now compromise much of the discussion of resource egalitarianism in the broader philosophical literature.² Thus, it is important to disentangle the various strains of Dworkin's argument, in order to see precisely what the auction contributes to the overall resource-egalitarian framework. The auction is, in fact, an oblique application of a well-known economic result known as the Second Fundamental Theorem of Welfare Economics.³ I would like to start out with a presentation of the envy-freeness standard that shows how the idea can be applied to simple cases without an auction mechanism. I then turn to the Second Fundamental Theorem, explaining briefly how it works and what it contributes to the overall resource-egalitarian picture. Using these two elements, I then go through Dworkin's application of the resource auction, in order to show what is right and wrong in his interpretation of it.

1. The envy-freeness standard

The idea of using the absence of envy as a criterion of equality has been discovered and rediscovered so many times that one is tempted to think it must be correct.⁴ The basic principle is stated with admirable clarity in Hobbes's *Leviathan*. Hobbes, it will be recalled, wanted to show that people in the state of nature were roughly equal in both physical and intellectual endowment. The level of variation in physical ability could be ascertained quite easily through inspection. Intellectual ability, on the other hand, presented something of a measurement problem (back in the dark ages, before the dawn of standardized testing).

Hobbes argued, therefore, that the most reliable indication that people were equal in their intellectual abilities is the tendency everyone has to think that they are smarter than everyone else. As Hobbes puts it, 'there is not ordinarily a greater sign of the equal distribution of any thing, than that every man is contented with his share'.⁵

One can see how the same features that recommended this principle to Hobbes, at this particular juncture of his argument, would also endear it to contemporary philosophers. There is, in contemporary discussions, both an acute awareness of the diversity of preferences and life projects, and a commitment to develop principles of justice that are neutral with respect to these differences. Thus, it is no longer persuasive to argue that we should simply equalize income, for example. Not all individuals value private goods equally, and so not everyone has the same use for income. But since these underlying preferences and projects are unobservable, it is difficult for an outsider to collect the information required to determine what would constitute an equal allocation. In this context, the envy-freeness standard recommends itself for much the same reason that it did for Hobbes. Since the information needed to achieve the equal allocation is private, the solution might be simply to let the individuals who have the information make their own choices under the right set of circumstances. When this exercise generates a set of compossible choices, one can then say that the allocation is equal.

To grasp this idea on a more intuitive level, consider a very mundane example. Imagine the situation of an executor of an estate who is instructed to divide up the deceased's possessions in such a way as to give each heir an equal share. Suppose, further, that it is impossible to achieve this by mechanically dividing the property up into identical bundles (for example, the estate may contain indivisible items, such as a clock or a dinner service). However, equalizing the bundles by determining the appraised monetary value of these items is also unsatisfactory, since it incorporates inessential factors into the deliberations (namely, how much people in society at large want the items) and excludes essential ones (namely, the sentimental value to the heirs).

One way of resolving the problem is to use the following procedure. The executor asks the heirs some preliminary questions about what each one wants, then divides up the estate's goods into a set of bundles. If there are three beneficiaries, the executor will make three bundles, and so on. Each beneficiary then writes down on a piece of paper which bundle he or she wants, and submits this request to the executor. If the set of requests is compossible (that is, each beneficiary requests a different bundle, and so all the requests can be satisfied simultaneously), then the bundles are distributed, each bundle given to the person who requested it. However, if any two heirs request the same bundle, then the goods are not distributed. The executor consults the preference schedules again, reorganizes 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.

This procedure is one way of generating an envy-free allocation. Once the allocation is complete, no one will envy anyone else's share. Of course, it is not the fact that no one 'feels envious' in the end that 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 they get under the proposed arrangement better than anyone else's. As a result, no one has any incentive to reject the arrangement. Thus, the envy-freeness principle is very closely tied to the very simple and attractive idea that a just cooperative arrangement must be one that is acceptable to all.

In the economic literature, this concept of envy-freeness is often used to define the following two distinct principles, one stronger than the other:

1. Fairness: a distribution is fair if, and only if, no individual prefers the bundle that is received by any other individual to his or her own.
2. 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.⁶

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 re-arrangements of the bundles. Under superfairness, each 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 to create identical bundles. (In cases in which some goods are indivisible, it is still possible to construct an equal allocation by creating a lottery, or by issuing shares, or by offering access rights, and so on.) Here there will be no envy, because every bundle is the same. This is often called the *equal* allocation. It is always 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). Since everyone will prefer more of that good to less, any bundle that contained more would be preferred by everyone. Similarly, if people have identical preferences, then only fair allocations will be available, not superfair ones. (Although these fair allocations need not be equal. 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, the existence of superfair allocations is a consequence of the heterogeneity of both goods and preferences.

There is one problem with the superfairness principle that should be noted. It does not pick out a unique solution. Furthermore, individuals will not be indifferent among the elements of the set of available superfair solutions. Imagine, for example, the division of an estate between two heirs, each of whom has a set of items that he or she is very strongly attached to – so strongly that it generates an

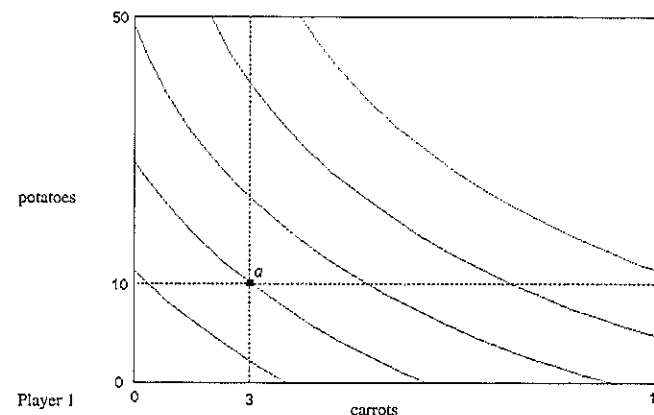


Figure 1 Indifference curve map

automatic preference for any bundle in which it is contained. If X is the set of items that the first heir strongly wants, Y is the set that the second strongly wants, and Z is the rest of the estate, then $(X \cup Z, Y)$ and $(X, Z \cup Y)$ are both superfair divisions. Prima facie this would suggest that it is of no consequence, from the standpoint of equality, how Z is to be allocated and that solutions in which it is all given to one person are perfectly fine. But this seems arbitrary, perhaps even unfair. In any case, it suggests that the envy-freeness standard needs to be supplemented with some additional principles in order to provide a persuasive theory of justice.

2. A graphical representation

There is a very simple graphical representation available of the set of superfair allocations which helps to show why this indeterminacy arises. However, it uses an Edgeworth box diagram, a mode of representation that strikes many people as unintuitive when they first encounter it. It is, nevertheless, worth the effort.

As we have seen, in order for there to be superfair solutions to an allocation problem, there must be at least two people and at least two goods to be distributed between them. The simplest way of representing an individual's preferences over two goods is with a set of indifference curves, as in Figure 1. For any point on such a curve, the agent is indifferent between that bundle and any other one that lies on the same curve; she will prefer any bundle that lies on a higher curve; and she will 'disprefer' any bundle that lies on a lower curve.

Once such a diagram is drawn, we can then represent the set of allocations that will make this individual envious. Assume that there are 10 carrots and 50 pota-

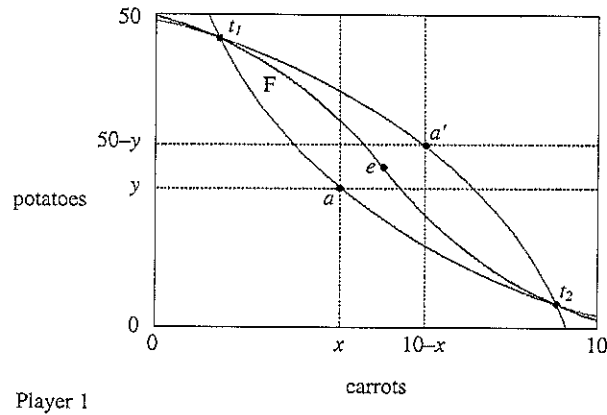


Figure 2 Detecting envy

atoes to be divided up between two individuals. Under any allocation, the amount that Player 1 gets (x, y) will lie upon a particular indifference curve. To see whether this allocation makes Player 1 envious, all we need to do is plot out the amount that the *other* player gets under this allocation $(10 - x, 50 - y)$, and see whether this point lies on a higher or a lower indifference curve than (x, y) . Figure 2 shows an example of this, with a being the amount that Player 1 will get under the proposed allocation and a' being the amount that Player 2 will get. We can see clearly that a will induce envy, since a' is on a higher indifference curve than a . We can check for envy with all the other points on the indifference curve that a is on just by plotting out what each of these allocations gives to the other player. This will give us a second curve that is a precise mirror image of the first. We can determine by inspection that almost every point on this indifference curve is one that will induce envy. It is only the set of points that lie on the curve beyond the two points where it intersects its mirror image (that is, above t_1 and below t_2) that do not produce envy. At either of these two points, the individual is indifferent between his allocation and the other player's.

Thus each indifference curve that intersects its own mirror image gives us two threshold points, which represent two allocations at which the individual ceases to prefer the other player's bundle. The entire set of such points gives us the 'fairness threshold' for that individual. Every allocation that lies on the inside of this threshold generates envy, while every point that lies beyond it does not. The fairness threshold for Player 1 is shown as line F in Figure 2. This line will always intersect the equal division point, shown here as e . The fact that it is convex above this point and concave below reflects the fact that this individual likes

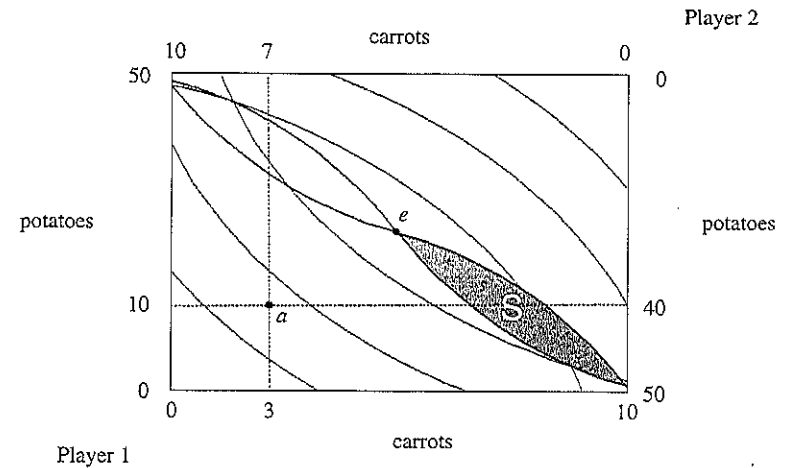


Figure 3 Superfair points

carrots more than potatoes, and so is more likely to be envious when her bundle contains a greater balance of potatoes than carrots.

Suppose that instead of representing one individual's preferences over various bundles of carrots and potatoes, we want to represent the preferences of two individuals over a division of carrots and potatoes between them. An Edgeworth box diagram allows us to do this in a very economical way. The diagram is constructed by taking two standard indifference curve maps, then 'flipping' one of them over, so that instead of having its origin at the south-west corner of the box, the origin will be at the north-east corner. By superimposing these two, as in Figure 3, it is possible to see both players' preferences over an allocation simultaneously. Thus, Figure 3 shows the division of 10 carrots and 50 potatoes between two individuals. For example, at point a Player 1 receives 3 carrots (shown on the lower axis), while Player 2 receives 7 (shown on the upper axis); Player 1 receives 10 potatoes (shown on the left axis), while Player 2 receives 40 (shown on the right axis).

Player 1's indifference curves are the ones that are concave to the south-west corner, while Player 2's indifference curves are convex to the south-west corner (and, of course, concave to her origin in the north-east corner). The set of superfair points will be those that are beyond both players' fairness thresholds. In Figure 3, this is the set of points in the region S. The reason that there are such points here is that the players have different, complementary tastes. Player 1 likes carrots more than potatoes, while Player 2 likes potatoes more than carrots. Thus,

the allocations to the south-east of the equal division, which give Player 1 relatively more carrots and Player 2 relatively more potatoes, contain a region of superfair points. If the two players had exactly the same preferences, then the two fairness frontiers would land on top of one another, and there would be no superfair solutions, only fair ones.

3. The Second Fundamental Theorem

The graphical representation presented here shows quite clearly why the envy-freeness principle does not provide a unique solution to the distribution problem, and thus why it will need to be supplemented with some other principle in order to provide a comprehensive theory of justice. Economists have traditionally regarded the Pareto-efficiency standard as the most obvious candidate for such a supplementary principle. This principle ranks one allocation superior to another if it is preferred by at least one individual and not 'dispreferred' by any other. The outcome is Pareto optimal when the limit of such improvements has been reached.

There is in general no reason to believe that an envy-free allocation will be Pareto optimal. Consider, for example, the estate-distribution problem outlined above, but one in which the executor decides to divide up Z into two sub-bundles, Z_1 and Z_2 . By hypothesis, both $(X \cup Z_1, Y \cup Z_2)$ and $(X \cup Z_2, Y \cup Z_1)$ will be superfair. But if one person prefers Z_1 to Z_2 while the other prefers Z_2 to Z_1 , then one of these allocations will be Pareto inferior to the other.

The set of Pareto optimal points in our allocation problem can also be represented quite economically using the same type of Edgeworth box diagram. Any particular allocation will fall on one of Player 1's indifference curves and on one of Player 2's indifference curves. That point will be Pareto optimal if it is impossible to move to a higher indifference curve for one player without moving to a lower indifference curve for the other. Thus, in Figure 4, point a is not Pareto optimal, because it is possible to pick another point, a' , which Player 2 likes just as much, but which is on a much higher indifference curve for Player 1. An allocation will be Pareto optimal only if it is the sole point at which the two players' indifference curves intersect. The set of such points is known as the contract curve, and is shown as line C in Figure 4.

This representation makes it very easy to see which superfair allocations will be Pareto optimal. It is the set of points on the contract curve C that lie within the region of superfair points S . This is shown in Figure 5. Again, if the two players have identical preferences, the contract curve will run through the equal division point. If the players have different preferences, we can rest assured that the contract curve will intersect S , because the points on the other side of e will be ones where both players envy each other's bundle. If they both envy each other, then the allocation cannot be Pareto optimal. They can simply trade bundles, which will take them to one of the points in S .

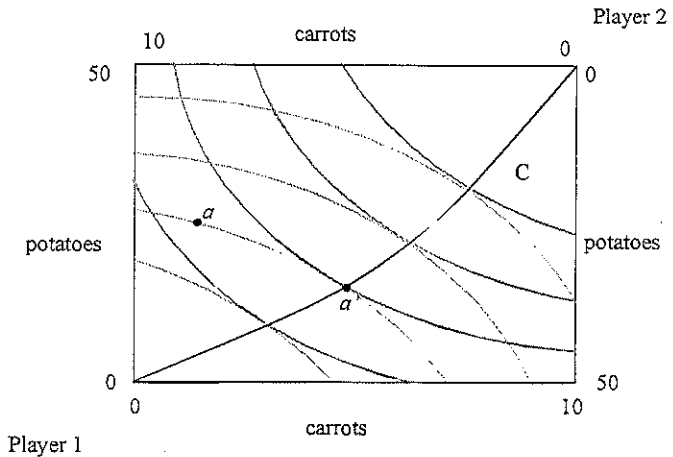


Figure 4 The contract curve

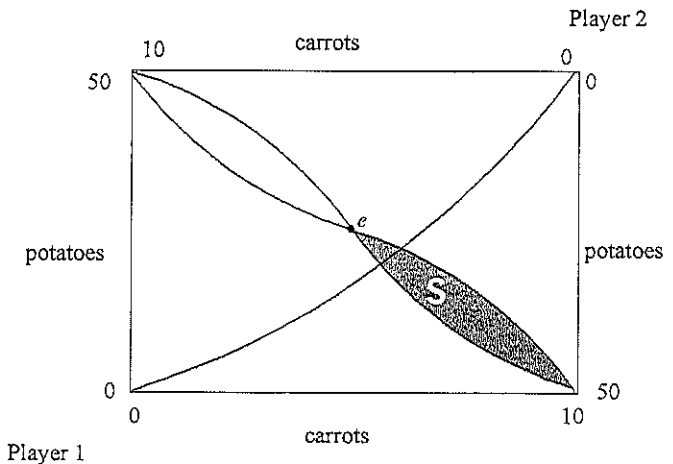


Figure 5 Superfair Pareto-optimal allocations

This observation leads naturally to a more general question. If the parties receive an allocation that is not superfair and Pareto optimal, under what circumstances will they be able to trade their way to such an outcome? More specifically, under what conditions will trading lead from an allocation that is not superfair to one that is? Also, will trading that starts at a superfair allocation necessarily lead to an outcome that is also superfair?

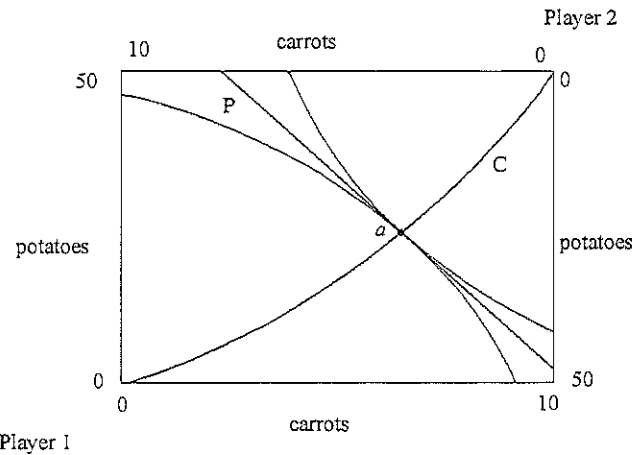


Figure 6 Second Fundamental Theorem

It is in order to answer these questions that we must appeal to the Second Fundamental Theorem of Welfare Economics. This theorem is, in effect, the converse of the better-known First Fundamental Theorem, which shows that the outcome of a perfectly competitive market will be Pareto optimal (also sometimes known as ‘the Invisible Hand Theorem’, with reference to Adam Smith⁷). The second theorem shows that, under certain assumptions, every Pareto optimal allocation can be the equilibrium outcome of a competitive market, given some initial allocation of goods and resources. This shows that there is no conflict in principle between equality and the exchange of goods. Whatever outcome on the contract curve the principle of equality favors, it is possible to construct a market that will generate precisely that outcome.

The second theorem also has a very simple graphical representation, shown in Figure 6. For any point on the contract curve C, such as *a*, one can draw a line that is tangent to both players’ indifference curves at that point. The point *a* is, then, the competitive equilibrium of a market that starts with any of the allocations on this line P. Of course, the characterization of the ‘market’ here is subject to the usual idealizing conditions: there must be no externalities, symmetrical information, no satiation, everyone must be a price-taker, and so forth.

Putting together the second theorem with the characterization of the set of superfair points suggests one very elegant way of solving the indeterminacy problem. As we have seen, even combining the envy-freeness standard with the Pareto-optimality principle still leaves us, in most cases, with a whole set of possible solutions and no way of privileging just one. However, we have seen that the equal solution *e* is always available and always unique. The problem with

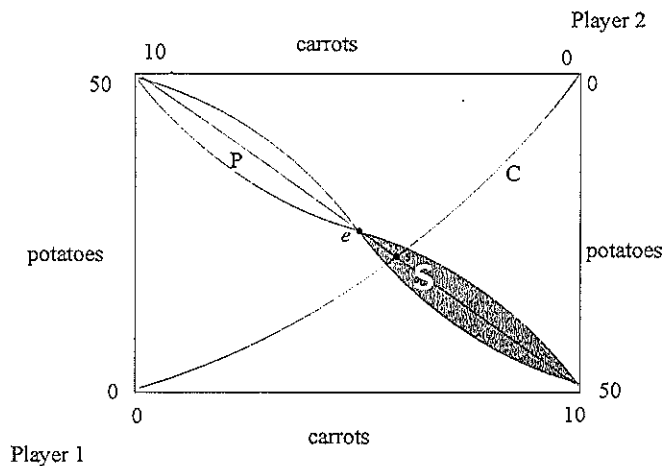


Figure 7 The efficient equal allocation

it, in the normal run of cases, is that it is inefficient (that is, not on the contract curve). So if one starts from this allocation, the players will generally be unhappy and refuse to stick with it. The second theorem, however, allows us to show that there will always be a point on the contract curve that ‘corresponds’ to it. This equilibrium point is shown as *s* in Figure 7. Call this the *efficient equal allocation*.

It is easy to show that, in cases in which the contract curve does not run through *e*, the competitive equilibrium of a market economy that starts with *e* as its initial allocation will also be superfair. This is because P is a price line – starting at any point on this line, in order to move to some other point on the line, the number of carrots that one must give up in order to get a potato, or vice versa, is the same. Thus, the bundles that Player 1 and Player 2 have at point *s* are both bundles that the other could have bought, had they preferred them. So one can show quite easily by *reductio* that point *s* must be superfair. More generally, one can show that trading under conditions of perfect competition will always preserve envy-freeness. If the initial allocation is envy-free, then the competitive equilibrium corresponding to that initial point will also be envy-free.⁸

This strategy for picking out a solution has a number of attractive properties. The basic problem with the envy-freeness standard, as shown in Section 1, is that some superfair allocations seem to be more ‘fair’ than others. However, the equal solution is unimpeachable in this respect, since it gives each person exactly the same amount. The only problem with it is that it is inefficient. So picking out the point on the contract curve that corresponds to the equal allocation solves both problems.

The other attractive feature of this solution is that it provides a way to engage in 'second-best' reasoning about equality. One unfortunate characteristic of the envy-freeness standard is that it simply partitions the set of allocations into those that create envy and those that do not. Among those that do, there is no way to determine how much envy they create, and so no way to rank them as more or less equal.⁹ Since it will be impossible, for the most part, to achieve envy-free allocations in the real world, this is a fairly serious problem. However, it is possible with a few modifications to rank allocations as more or less equal. We can begin by drawing a diagonal line from one player's origin to the other's, intersecting point e on the way. Allocations on this line can then be ranked as more or less equal, depending upon how far away they are from the equal allocation. All of the points on the contract curve can then be ranked by finding the competitive equilibrium corresponding to each point on the diagonal, and using the ranking of points on the diagonal to rank the points on the contract curve.¹⁰

4. Dworkin's argument

With this analysis in hand, we can now turn to Dworkin's presentation of his auction. He introduces the auction, as we might expect, in order to handle the indeterminacy of the envy-freeness standard. The shipwreck survivors, having agreed to divide up all of the resources on the island equally (where 'no division of resources is an equal division if, once the division is complete, any immigrant would prefer someone else's bundle of resources to his own bundle' (SV: 67)), must settle upon some mechanism for doing so. Initially, Dworkin considers the possibility of a mechanical division of the resources into equal bundles, or else the use of a procedure similar to that of the estate executor (outlined above). But he argues that such a division, even if physically possible, will not eliminate a possible 'piece of unfairness':

For the combination of resources which composes each bundle the divider creates will favor some tastes over others, compared with different combinations that he might have composed. That is, different sets of n bundles might be created by trial and error, each of which would pass the envy test, so that for any such set that the divider chooses, someone will prefer that [the divider] had chosen a different set, even though that person would not prefer a different bundle within that set. (SV: 67)

One can see what Dworkin is getting at here, even though his statement of the problem is not quite right. What Dworkin is describing is simply the fact that the envy-freeness standard picks out an entire region of points, and that, even though, for any given allocation, no one prefers anyone else's bundle *within* that allocation, at least one person will prefer the bundle that he would receive *under some other envy-free allocation*. But this by itself cannot be the problem. The fact that there is a region of envy-free allocations means that, no matter which allocation gets chosen, there will always be someone who prefers some other allocation.

Nothing (including Dworkin's proposed solution) can do anything to eliminate this. The problem therefore cannot be simply that 'someone will prefer that [the divider] had chosen a different set'. The true problem is raised only later, when Dworkin states that 'if any more complex division can be found that will satisfy [the envy test], many such might be found, so that the choice among these will be arbitrary' (SV: 68).

Stated more precisely, the problem is as follows. We have good reason to choose an envy-free allocation over one that induces envy. But given the set of envy-free allocations, what reason do we have to choose one rather than another? The mere fact that an 'executor' happens to have settled upon one such allocation cannot be a good reason for choosing it. That is arbitrary, and hence unfair to those who receive less under this allocation than under some other.

Luckily, says Dworkin, after this statement of the problem, 'the same solution will by now have occurred to all readers. The divider needs some form of auction or other market procedure in order to respond to these problems' (SV: 68). This moves far too quickly, since there are a variety of ways that we might solve the indeterminacy problem without bringing in an auction. For instance, we might choose to treat the region of superfair allocations as a second division problem, and apply the envy-freeness standard to it.¹¹ Recursive application of the envy-freeness standard will eventually converge toward a limit point. This solves the problem, even though the solution has its weaknesses. The point is simply that an auction does not present itself as the only possible solution to the indeterminacy problem.

In any case, having settled upon an auction as the appropriate sort of procedure for privileging one envy-free solution, Dworkin goes on to characterize the auction procedure in the following way:

Suppose the divider hands each of the immigrants an equal and large number of clamshells, which are sufficiently numerous and in themselves valued by no one, to use as counters in a market of the following sort. Each distinct item on the island (not including the immigrants themselves) is listed as a lot to be sold . . . The auctioneer then proposes a set of prices for each lot and discovers whether that set of prices clears all markets, that is, whether there is only one purchaser at that price and all lots are sold. If not, then the auctioneer adjusts his prices until he reaches a set that does clear the markets.

Dworkin then concludes, correctly, that the allocation which results from such an auction will be envy-free, since 'no one will envy another's set of purchases because, by hypothesis, he could have purchased that bundle with his clamshells instead of his own bundle' (SV: 68).

The problem with this construction is a lack of clarity about what sort of work the auction is doing. It should be recognized that what Dworkin is proposing here as a solution to the distribution problem is identical to the 'efficient equal' solution sketched out at the end of the previous section.¹² This is obscured by

Dworkin's use of the clamshells as an expository device. The initial distribution, in which each survivor receives an equal number of clamshells, is equivalent to the equal allocation of resources (point e). To see why, consider how the auction would work in the simpler two-person case, in which the problem is to divide up 10 carrots and 50 potatoes. Suppose that 100 clamshells are to be distributed between the two. Instead of giving out clamshells, we could also give out little pieces of paper that say: 'this entitles the bearer to a tenth of a carrot, and half a potato' (since receiving one clam shell is the same thing as receiving a hundredth share of all the resources). We could then conduct an auction along the same lines as Dworkin's, with identical results. Thus, Dworkin's use of clamshells runs the risk of generating a 'money illusion' – concealing the fact that his auction takes the equal allocation as its point of departure.

Once we have seen through this money illusion, we can then ask precisely what the auction accomplishes. It does not serve to create equality, since the initial allocation is already perfectly equal and envy-free. The problem with the initial allocation is simply that it is inefficient. Thus, the auction takes the initial envy-free allocation and transforms it into an efficient envy-free allocation (by taking the survivors from the equal allocation to the corresponding point on the contract curve).¹³ The fact that it can do so is a relatively trivial implication of the Second Fundamental Theorem, since the auction is simply a stand-in for a perfectly competitive market, and trading under conditions of perfect competition preserves envy-freeness.¹⁴

If one eliminated the auction entirely, and simply left the survivors with the equal allocation, none of them could complain that they had been treated unequally. So the auction does not appear to be accomplishing much on the equality front. Dworkin, however, believes that there is a deeper connection between the auction and the ideal of equality. He argues that the division of the resources into bundles by an executor is different from the division produced through an auction because through the auction 'each person played, through his purchases against an initially equal stock of counters, an equal role in determining the set of bundles actually chosen' (SV: 68). This is what eliminates the 'arbitrariness' associated with an executor's division of resources.

Unfortunately, Dworkin does not explain what it means to play an 'equal role' in this context.¹⁵ Clearly, the auction does nothing to influence the *form* that the resources take. 'If plover's eggs and old claret were the only resources to auction, then the person who hated these would be as badly off as in our earlier example [of the mechanical division] . . . but he could not complain that the division of the actual resources they found was unequal' (SV: 69). Questions concerning the 'baseline' are dealt with by the principle of abstraction (introduced later in Dworkin's work on liberty), and not by the auction.¹⁶ Thus, the notion of 'an equal role in determining the set of bundles' can only refer to the bidding power each individual has, within the auction, that allows him or her to influence which resources will be lumped together to form a bundle. But since the bidding power

of each individual is determined by the number of clamshells that each is given at the beginning of the auction, any equality of bidding power must be a function of the equality of the initial distribution. In other words, the auction does nothing to guarantee everyone an 'equal role', it is entirely the choice of e as the initial allocation that equalizes roles. Thus, the outcome that results from the auction is neither more nor less equal than the initial allocation. It is simply more efficient.

5. Problems with Dworkin's argument

These reflections do not in any way impugn the integrity of the resource-egalitarian ideal. The problem is simply that Dworkin overestimates how central the auction is to this ideal, because he fails to see that the initial handout of clamshells amounts to an equal (and envy-free) allocation of the island's resources. Thus, he is led to state, for instance, that 'an equal division of resources presupposes an economic market of some form, mainly as an analytical device but also, to a certain extent, as an actual political institution' (SV: 66). As we have seen, the market has nothing to do with the equal division of resources. It is easy to create an equal division without a market, just by giving everyone an identical bundle. The tricky part is to create an efficient division – it is here, and only here, that we require the services of the market.

Dworkin states that the market has erroneously 'come to be regarded as the enemy of equality, largely because the forms of economic market systems developed and enforced in industrial countries have permitted and indeed encouraged vast inequality in property' (SV: 66). This is correct. As the Second Fundamental Theorem shows, inequality of outcome is a contingent, not a necessary, feature of the market pattern of organization. However, what the second theorem shows is that markets are transparent to issues of equality or inequality. If the input is unequal, the output will be unequal, and if the input is equal, the output will be too. Thus, Dworkin overstates his case when he claims that: 'the idea of an economic market, as a device for setting prices for a vast variety of goods and services, must be at the center of any attractive theoretical development of equality of resources' (SV: 66). What the second theorem shows is that while there are no egalitarian arguments *against* the market, neither are there any egalitarian arguments *for* the market. This is probably a good thing, since an egalitarian argument for the market would be easy to caricature and dismiss, as Elizabeth Anderson has demonstrated.¹⁷ Efficiency arguments are far more compelling.

There is another line of thinking in Dworkin's work which suggests that the resource-egalitarian ideal requires markets because property rights force individuals to take account of the costs that their choices impose upon others. Thus, Dworkin sometimes claims that showing 'equal concern' for all persons requires a market, because the market forces 'each to take responsibility for the true costs of his own choices'.¹⁸ When Bill eats a carrot, that is one less carrot available for the rest of humanity. In the absence of markets, Bill need not take into account

the extent to which other people might want the carrot, prior to his eating it. However, with the institution of the market, the needs of others will be reflected in the price at which carrots are selling. This may increase the opportunity cost of Bill's choice. Instead of simply eating a carrot, he may be eating a carrot worth \$5 – a reflection of the fact that other people desperately want it. He will therefore be foregoing all of the other consumption bundles that \$5 could purchase.

There is, of course, nothing wrong with the moral intuition underlying this argument for the market. The problem is that the intuition stems from our commitment to efficiency, not equality. Not only Dworkin, but many of his commentators have been misled in this regard.¹⁹ The example of Bill's carrot is just a concrete example of the way that markets bring about an alignment of private and social costs. As Dworkin puts it, markets fix 'the value of any transferable resources one person has as the value others forgo by his having it' (SV: 149). This is just another way of saying that property rights internalize externalities.²⁰ But the problem with externalities is that they generate waste and inefficiency, not inequality. If Ted wants Bill's carrot more than Bill does, and there is an unrealized exchange between them, this means that (relatively speaking) Bill must want Ted's potato more than Ted does. Thus, Bill's failure to take into consideration the cost to Ted of his choices in the absence of a market is necessarily matched by Ted's failure to take into consideration the cost to Bill of his own choices. This certainly constitutes a lack of *concern*, but it does not represent an affront to equality, because it is necessarily symmetric (otherwise it could not be corrected through market exchange).²¹ Thus, a society that imposes equal allocation as a pattern of distribution, then prohibits all exchange, nevertheless exhibits equal concern for both Bill and Ted. This concern takes the form of an *equal lack of concern* about their particular vegetable preferences. If that sounds odd, it is merely because Dworkin's phrase 'equal concern' smuggles in a commitment to efficiency in the form of the assumption that concern should be equalized at a high rather than a low level. (After all, if the primitive equality that Hobbes described in the state of nature in fact obtained, then that state would be perfectly envy free. Everyone would be equally unconcerned about everyone else, and everyone's life would be unenviably solitary, poor, nasty, brutish, and short.)

Thus, there are good reasons for wanting an institution such as the system of property rights that underlies Dworkin's auction. However, the reasons for it are all grounded in the attractiveness of efficiency as a normative ideal, not equality. This is, unfortunately, not an observation that Dworkin is well positioned to take advantage of, since he insists on excluding efficiency considerations from his articulation of the resource-egalitarian ideal. This is in contrast to someone like John Rawls, who believes that the core principles in a theory of justice will need to have efficiency considerations built in. (The difference principle states, in effect, when trade-offs between equality and efficiency are acceptable.²²) Dworkin, however, insists upon a peculiar sort of normative monism, according to which the principle of equality must be 'an independent and powerful politi-

cal ideal' (SV: 109) which should not be compromised or diluted by other principles.²³ It is perhaps this commitment to 'equality monism' that leads him to obscure the nature of the contribution that the auction makes to his scheme. If Dworkin were consistent in his insistence upon a monistic conception of equality, he would have to opt for recursive application of the envy-freeness standard, instead of the auction, as a way of picking out a unique solution.

Part of the reason that Dworkin insists so much upon the separation of equality and efficiency is that he often flouts the distinction between the Pareto-efficiency standard and the utilitarian injunction to maximize average utility (SV: 56). The difference is quite significant, however, since the Pareto conception of efficiency is perfectly compatible with equality (as the second theorem, along with Dworkin's own auction, shows), while the injunction to maximize average utility often generates a conflict with equality. So, even though Dworkin finds the idea of an auction attractive and essential, he feels that he cannot grant that it responds to an efficiency concern. Thus, he represents the auction as though it were making some contribution toward the equality of the outcome. He even states, at one point, that under the right sort of circumstances, 'efficiency simply is fairness, at least as fairness is conceived under equality of resources' (SV: 84). This is deeply muddled.

Of course, Dworkin does not intend to ignore efficiency considerations completely. He simply wants to formulate a principle of equality that is uncontaminated by efficiency considerations, or, as he puts it, to 'protect the boundaries of the concept of equality from confusion with other concepts' (SV: 109). But he anticipates that efficiency will enter into the picture at a later stage, as a concession to the 'real-world' difficulties that one is likely to encounter when it comes time to implement egalitarian arrangements. Thus, he acknowledges the 'familiar idea in political theory that a just society will make some compromise between efficiency and distribution. It will sometimes tolerate less than perfect equality in order to improve average utility' (SV: 54).²⁴

The idea that we might start out with a purely normative theory, then make various concessions to practical concerns as we seek to establish functional economic and political institutions is, however, subject to some notorious difficulties. In particular, because Dworkin's use of the auction mechanism is simply an application of the Second Fundamental Theorem, his project is vulnerable to all of the objections that have been raised, over the years, to classic welfare economics. In particular, the idea that a market mechanism might be used to implement an egalitarian distribution of wealth is a very old idea, and has been central to debates over 'market socialism' since the 1940s.²⁵ Dworkin's framework is subject to all of the objections that this project has attracted over the years.²⁶

In order to see the seriousness of these objections, it is helpful to recall just how powerful the idealizations are underlying the conception of a perfectly competitive market (or a Walrasian auction). Simple trading will not, in general, preserve envy-freeness. Suppose that Bill and John exchange carrots for potatoes,

but then Fred comes along, hoping to do the same. Unfortunately for him, Bill has already bought all the potatoes he wants from John. This leaves Fred envying John, who had the chance to unload all of his surplus potatoes. The problem here is that exchange has occurred at a price level that does not clear the market. Trade will only preserve envy-freeness if it occurs at exactly that price. But market-clearing prices will emerge only under the highly restrictive set of conditions that characterize perfect competition (known as the Pareto conditions).

What do these conditions include? They include, for instance, a complete set of property rights, so that there will be no positive or negative externalities. (This is why Dworkin states that *all* resources, except 'the immigrants themselves', are to be divided up into lots.) This set of property rights extends far beyond the set of rights that are legally recognized in our society. They would have to cover everything that is a public good in our society – not just land, but also air, water, security, green space, and so on. But it goes much further. Imagine any sort of externality – such as making a lot of noise that keeps the neighbors awake. In a perfectly competitive market, there would have to be a property right that fully internalized the cost that this action imposes upon others. Without such rights, there would be the usual overproduction of goods associated with negative externalities (noise, smoke, odor, and so on) and the underproduction of goods associated with positive externalities (construction of aesthetically pleasing huts, killing of poisonous snakes, and so on).

Most of these types of property rights do not exist in our society, because the transaction costs associated with their enforcement would be prohibitive. Many people have thought, however, that even if real-world markets can never satisfy these ideal conditions, they can at least come fairly close. Dworkin assumes this, claiming that 'under certain (perhaps very limited) circumstances, when the conditions for an equal auction are at least roughly met, then an actual auction might be the best means of reaching or preserving equality of resources in the real world'. He goes on to state that the 'economic markets of many countries can be interpreted, even as they stand, as forms of auctions' (SV: 72).

This is a very standard sort of argument. The ideal of a perfectly competitive market, according to this view, is like the idea of a frictionless plane in physics. Even though we are never in the perfectly ideal state, the more closely the world resembles the ideal state, the more closely the results will approximate those predicted by ideal theory. On the basis of this analogy, we may be tempted to conclude that if perfect competition generates perfect efficiency, then near-perfect competition should generate something as close as possible to perfect efficiency. So a real-world, resource-egalitarian auction should generate something as close as possible to an egalitarian allocation.

The problem for this view arises from what is known as the 'general theory of the second best'.²⁷ This theorem shows that, in a situation in which one of the Pareto conditions is violated, respect for all of the other Pareto conditions will generate an outcome that is less efficient than some other outcome that could be

obtained by violating one or more of the other conditions. In other words, while perfect competition generates a perfectly efficient outcome, a situation that is as close as possible to perfect competition will not generate an outcome that is as close as possible to perfect efficiency. Thus, the analogy between perfect competition and a frictionless plane is a fallacy.

The primary consequence of the second-best theorem is that it deprives both the First and the Second Fundamental Theorems of Welfare Economics of any real-world implications. Moreover, because Dworkin's auction is an application of the Second Fundamental Theorem, it is subject to all of the limitations imposed by the second-best theorem. The relationship between equality and efficiency that it reveals, and that the auction realizes, obtains only at the level of ideal theory. The thought experiment cannot be used to derive any conclusions about the way that actual markets work in the world or how actual resource-egalitarian allocations can be achieved.

It is important to emphasize the nature of the problem here. It is not that an auction conducted under less-than-ideal conditions will not produce an efficient outcome. That is a problem, but it is not necessarily one that Dworkin need worry about. The problem is that an auction that does not generate a perfectly efficient outcome will not preserve envy-freeness. Thus, even if one begins with a state of perfect equality, a real-world auction is almost guaranteed to destroy it. This is an issue that Dworkin cannot ignore.

Lastly, it is worth noting that Dworkin's reliance upon the traditional welfare economics assumptions that underlie the Second Fundamental Theorem creates a rather unexpected opportunity for a welfarist counterattack. Economists have traditionally defined positive and negative externalities in terms of the effects that certain transactions have upon the *welfare* of third parties. As a result, the idea of a system of property rights that internalizes all externalities is essentially equivalent to a system in which anyone who wants to reduce anyone else's welfare must pay him or her to do so. But this could just as well be achieved through a system of property rights that gave each individual ownership of his or her own welfare. Thus, the difference between equalizing resources and equalizing welfare turns out to be less than clear. (In a sense, Dworkin's argument trades upon an equivocation in the sense of the term 'resource', between what we in our society call a resource and what would have to count as a resource under a system of property rights that internalized all externalities.) In the end, Dworkin's resource egalitarianism may appear persuasive only because, under the set of idealizing assumptions that it relies upon, it is analytically equivalent to one form of welfarism.

6. Conclusion

My central contention in this article has been that Dworkin misunderstands the contribution that his auction makes to his resource-egalitarian scheme. This leads

him to make two illegitimate inferences concerning the status of the market. First, it leads him to think that there is some internal connection between the institution of the market and the resource-egalitarian ideal.²⁸ This is incorrect. Markets promote efficiency; they are at best neutral with respect to equality. Second, it leads him to overlook results in economic theory which show that *in principle* conclusions derived at this level of idealization have no normative implications for real-world institutions. General equilibrium theory simply cannot be used as a platform for the development of a theory of justice.

My discussion has been focused entirely upon Dworkin's preliminary characterization of his auction, involving the story of the shipwreck survivors. Dworkin goes on to make the construction significantly more complex by insisting that the allocation be not only statically envy free, but also dynamically envy free (roughly speaking). Since the initial auction occurs under conditions of uncertainty about the future, Dworkin adds in a set of complementary insurance markets. He then argues, for reasons somewhat tangential to the concerns discussed here, that there should be a set of hypothetical insurance markets to compensate individuals for various forms of undeserved bad luck that they have been unable to insure themselves against. The resulting construction is extremely complex, and also deeply problematic.²⁹

Such extensions of the project are, however, impossible to evaluate without a solid understanding of how the basic auction works. Insurance systems have not only important effects upon the distribution of wealth, they also confer significant efficiency gains upon the parties involved. These two components must be separated out in order to assess the relevance of such arrangements to the envy-freeness standard as a conception of equality. Once these two strands are separated out, however, my suspicion is that the primary rationale for social-insurance-type arrangements will be grounded in their efficiency properties, in the same way that the primary justification for the market rests with its efficiency properties. Dworkin's attempt to defend these institutional arrangements through reference to their supposedly egalitarian character generates a series of extremely weak arguments, and runs the risk of trading upon a straightforward conceptual confusion.

My hope is that the discussion in this article will provide some of the theoretical tools needed to work through these difficulties. Both the resource-egalitarian framework and the envy-freeness standard have, in my view, considerable merit as components of a broader theory of equality. Dworkin's auction, unfortunately, does not do justice to either. His presentation unhelpfully saddles the envy-freeness idea with a set of inessential commitments. Separating out the strands of the argument, in order to see precisely what service the auction mechanism performs, is the first step toward a more adequate evaluation of the merits of envy-freeness as a principle of equality.

notes

1. See Ronald Dworkin, 'Equality of Welfare' and 'Equality of Resources', in *Sovereign Virtue* (Cambridge, MA: Harvard University Press, 2001).
2. For example, Will Kymlicka's use of Dworkin's auction thought experiment to justify minority rights is based on a flawed interpretation, as it assumes the presence of uninternalized 'network externalities' associated with language and cultural practices. See Will Kymlicka, *Liberalism, Community and Culture* (Oxford: Clarendon, 1989). Similarly, Elizabeth Anderson's recent reappraisal of the 'point' of egalitarianism takes Dworkin's interpretation of the auction at face value, which allows her then to accuse resource egalitarianism of harboring a 'Poor Law' mentality. See Elizabeth Anderson, 'What is the Point of Equality?' *Ethics* 109 (1999): 287–337.
3. Economists will often describe Dworkin's egalitarian solution in a way that renders this explicit (for example, as 'the Walrasian equilibrium associated with an equal division endowment'). This sort of language is, unfortunately, opaque to the majority of political philosophers.
4. The following is an incomplete list: Duncan Foley, 'Resource Allocation in the Public Sector', *Yale Economic Essays* 7 (1967): 45–98; Hal Varian, 'Equity, Envy and Efficiency', *Journal of Economic Theory* 9 (1974): 63–91; William J. Baumol, *Superfairness* (Cambridge, MA: MIT Press, 1986); Serge-Christophe Kolm, *Justice and Equity*, trans. Harold F. See (Cambridge, MA: MIT Press, 1997).
5. Thomas Hobbes, *Leviathan*, edited by Richard Tuck (Cambridge: Cambridge University Press, 1991), p. 87.
6. Baumol, *Superfairness*, p. 19. The two concepts are comparable to weak Pareto superiority (which permits indifference among some) and strong Pareto superiority (which does not).
7. See Mark Blaug, *Economic Theory in Retrospect*, 4th edn. (Cambridge: Cambridge University Press, 1985), Ch. 13.
8. There is one noteworthy restriction to this claim. The example above is of a pure exchange economy, without production. Adding production into the mix complicates things, because it introduces a variety of problems pertaining to the distribution of leisure. See Elisha A. Pazner and David Schmeidler, 'A Difficulty in the Concept of Fairness', *Review of Economic Studies* 41 (1974): 441–3.
9. Note that Dworkin (SV: 72) claims, erroneously, that the envy-freeness standard alone can be used to rank distributions as more or less equal.
10. See Marc Fleurbaey and François Maniquet, 'Fair Social Orderings' (forthcoming).
11. See Baumol, *Superfairness*, pp. 64–70.
12. This has been noted elsewhere. See, for example, John G. Bennett, 'Ethics and Markets', *Philosophy and Public Affairs* 14 (1985): 197.
13. Elsewhere, Dworkin describes the function of the auction as equalizing aggregate opportunity costs. See Ronald Dworkin, 'The Place of Liberty', in *Sovereign Virtue* (Cambridge, MA: Harvard University Press, 2001), p. 149. Again, he fails to notice that the auction does no such thing. Aggregate opportunity costs are already equalized by the initial allocation (since the bundles are identical, aggregate opportunity costs must be the same). The auction simply preserves the equality of these aggregate opportunity costs, while selecting a more efficient allocation.

14. Dworkin says: 'I mean to describe a Walrasian auction in which all productive resources are sold'. Walras famously demonstrated that under certain conditions an auction could achieve the same outcome as a perfectly competitive market. For a helpful non-technical overview, see Blaug, *Economic Theory in Retrospect*, Ch. 13.
15. The uncharitable conclusion, drawn by Hal Varian among others, is that Dworkin simply has no argument for privileging the equal, efficient outcome. See Hal Varian, 'Dworkin on Equality of Resources', *Economics and Philosophy* 1 (1985): 110–25. Varian notes that most economists 'added the additional requirement that an "equitable" allocation be Pareto efficient' as well as envy-free because they 'view Pareto efficiency as a desirable thing'. See *ibid.*, p. 112. Dworkin rejects this rationale, criticizes the initial equal allocation as 'arbitrary', but then fails to offer 'any justification' for using an auction as a way of resolving the arbitrariness. In the end, Varian says, 'Dworkin *feels* that the market mechanism is central to the notion of justice . . . but I don't think he has really *argued* for it'. See *ibid.*, p. 113.
16. See Dworkin, 'The Place of Liberty', *Sovereign Virtue*, pp. 150–1.
17. Anderson suggests that the market offers 'no aid to those it labels irresponsible, and humiliating aid to those it labels innately inferior. It gives us the cramped vision of the Poor Laws, where unfortunates breathe words of supplication and submit to the humiliating moral judgments of the state'. See Anderson, 'What is the Point of Equality?', p. 308.
18. Ronald Dworkin, 'Why Efficiency?' in *A Matter of Principle* (Cambridge, MA: Harvard University Press, 1985), p. 270. Elsewhere in the same volume, he says that the price system is needed to achieve equality because 'it fixes the true cost to the community of meeting [each individual's] own preferences for goods and activities'. See Ronald Dworkin, 'Liberalism', in *A Matter of Principle* (Cambridge, MA: Harvard University Press, 1985), p. 194.
19. Arthur Ripstein, for example, writes that the 'appeal of [Dworkin's] market-based account of equality was that the notion of opportunity cost appeared to provide a way of measuring the costs of each person's choices to *others*'. See Arthur Ripstein, *Equality, Responsibility and the Law* (Cambridge: Cambridge University Press, 1999), p. 284. Thus for Dworkin, the idea of opportunity costs reflects 'an interpretation of equality'. See *ibid.*, p. 280.
20. I am using the term 'externality' in the very broad sense pioneered by David Gauthier, *Morals by Agreement* (Oxford: Clarendon, 1986).
21. It is perhaps worth mentioning that there are plenty of welfarist arguments to show that when preferences differ, the equal allocation fails to respect equality. Precisely because preferences differ, each individual's endowment under the equal allocation will create opportunities for different (that is, unequal) levels of welfare. But Dworkin obviously cannot appeal to this intuition. Furthermore, the same welfarist objections to the equal allocation would apply against the efficient equal allocation that Dworkin favors.
22. This is sometimes overlooked. The difference principle favors perfect equality (according to some index of primary goods), *unless* it is possible to make Pareto-efficiency gains that increase inequality. When that occurs, it recommends choosing the Pareto improvement that maximizes the allocation of the worst-off. See John Rawls, *A Theory of Justice* (Belknap: Harvard University Press, 1971), p. 76. Rawls must assume that the question of trade-offs between the two principles will arise in the normative component of his theory in part because he does *not* adopt the idealizing framework of welfare economics that structures Dworkin's discussion.
23. See also Dworkin, 'Why Efficiency?'. The effects of Dworkin's strict partitioning of equality and efficiency concerns becomes clearly apparent in the way that he deals with the problem of expensive tastes. One of the difficulties with welfarism is that it offers to compensate fully individuals who acquire preferences that reduce the amount of welfare they can extract from a given bundle of resources. This generates a huge moral hazard problem, since it allows individuals to externalize fully the costs associated with the acquisition of such 'expensive' preferences. However, Dworkin argues that these considerations do not count against welfare egalitarianism, since they represent a mere efficiency concern. He insists that 'expensive tastes are embarrassing for the theory that equality means equality of welfare precisely because we believe that equality, considered in itself and apart from questions of efficiency, condemns rather than recommends compensating for deliberately cultivated expensive tastes'. See *ibid.*, p. 55. Thus, he falls back upon the much weaker argument that our reticence to indemnify expensive preferences reflects a prior tacit commitment to a non-welfarist conception of equality.
24. Again, one can see the very unhelpful confusion between efficiency and average-utility maximization. Note also the contrast to Rawls, who treats efficiency as a normative principle, not something that simply intervenes at the implementation stage.
25. Oskar Lange, *Essays on Economic Planning*, 2nd edn. (Bombay: Asian Publishing House, 1967); Oskar Lange and Fred M. Taylor, *On the Economic Theory of Socialism* (New York: McGraw Hill, 1964); Abba P. Lerner, *The Economics of Control* (New York: Macmillan, 1944).
26. Joseph Stiglitz, *Whither Socialism?* (Cambridge, MA: MIT Press, 1994).
27. Richard Lipsey and Kelvin Lancaster, 'The General Theory of the Second Best', *Review of Economic Studies* 24 (1956): 11–32.
28. See, for example, Dworkin's remarks at the top of p. 70, *Sovereign Virtue*.
29. See John Roemer, *Theories of Distributive Justice* (Cambridge, MA: Harvard University Press, 1996), p. 256.