

ECLECTIC DISTRIBUTIONAL ETHICS

By

John E. Roemer

March 2003

COWLES FOUNDATION DISCUSSION PAPER NO. 1408



COWLES FOUNDATION FOR RESEARCH IN ECONOMICS

YALE UNIVERSITY

Box 208281

New Haven, Connecticut 06520-8281

<http://cowles.econ.yale.edu/>

“Eclectic distributional ethics”

by

John E. Roemer¹

Yale University

Abstract: Utilitarians, egalitarians, prioritarrians, and sufficientarians each provide examples of situations demonstrating, often compellingly, that a sensible ethical observer must adopt their view and reject the others. We argue, to the contrary, that an attractive ethic is eclectic, in the sense of coinciding with these apparently different views in different regions of the space of social states.

¹ Departments of Political Science and Economics.

1. Introduction

My aim is to study the ethics of distribution in very simple societies, ones consisting of two people, whom I shall call A and B . There are at least four distributional ethics one might hold for these societies: egalitarian, prioritarian, utilitarian, or sufficientarian. I will explain these views precisely below. Advocates of each of these views often seem to have compelling arguments against the others, and this could be taken to imply some kind of ‘impossibility theorem’ for ethics. I will argue, on the contrary, that an attractive ethic requires that one hold *all* of these views in a restricted way: depending upon the problem, ethics may require one to be an egalitarian or a prioritarian, or something else. Moreover, I hope to show that this eclectic view is attractive, even within the narrow confines of a welfarist setting.

Welfarism is the view, first, that everything of value about a person’s life can be summed up in a number that measures his or her welfare, and that, second, a distributional ethic need only rank possible distributions of welfare, in a population, to be complete. As I here adopt a welfarist framework, I am not concerned with the ‘equality-of- what’ debate, which focuses upon what features of the human condition should be the objects of distributional concern. A *social state* in this paper will be an ordered pair of welfare levels, $u = (u_A, u_B)$, for A and B . The domain of conceivable social states is the set of all possible ordered pairs of non-negative numbers, called the non-negative orthant, and denoted \mathbf{R}_+^2 .

I identify four particular welfare levels for individuals:

the welfare level zero, associated with dying a horrible death (this is the lowest possible welfare level);

a welfare level b , associated with a comfortable death;

a welfare level b'' , associated with a death somewhat less than comfortable, but not horrible;

a welfare level b' , associated with a good life.

Clearly, $b' > b > b'' > 0$.

If a person has a welfare level greater than b , I say his life is *worthwhile*, or his life is worth living; if a person has welfare level of b' or greater, I say that she is *flourishing*. A person might sustain a welfare level of less than b , and be alive: he could be enduring a state of pain that is worse than having a comfortable death. Such occurrences, of course, are the justification of euthanasia.

In this paper, u_i (with a subscript) will always indicate the welfare level of person i , and u^j (with a superscript) will always indicate a particular social state, that is, an ordered pair of welfare levels.

2. Types of distributional ethic

A distributional ethic is a complete order on the set of possible social states, \mathbf{R}_+^2 . We denote the fact that state u^1 is ethically at least as desirable as the state u^2 in the usual way: $u^1 \succeq u^2$. The notation $u^1 \succ u^2$ is the *strict* preference relation.

It is convenient to represent preference orders by their *indifference curve maps*. An indifference curve is a set of social states that are ethically indifferent, according to a particular view. An indifference curve map is a set of indifference curves associated with

a particular view, with the property that every point in \mathbf{R}_+^2 lies on one and only one indifference curve in the map. An indifference curve map thus completely specifies an observer's ethical preferences.

I display four important types of preference order, by their indifference curve maps, in Figure 1.

[p figure 1 abcd about here]

Preferred social states always lie on 'higher' indifference curves.

To understand how indifference curve maps work, we first introduce the idea of a *social situation*. A social situation is a set of social states, which are available in a particular instance or problem. We assume that one of these states must be chosen. For example, the set T in figure 2, is a social situation. (In economics, a social situation is called a *utility possibilities set*.)

[Figure 2 here]

Suppose the social situation were T , and one's ethical preferences were given by the indifference map illustrated in Figure 2. Then the observer would choose the social state p , that state in T which lies on the highest indifference curve that intersects T .

We can now explain the names of the preferences in Figure 1. The preference relation in Figure 1a is called egalitarian for the following reason. Let T be a social situation, illustrated in Figure 3a, and let e be the Pareto efficient point at which the 45° ray intersects the boundary of T . e is the equal-welfare point that has the highest value of welfare, among all equal-welfare points in T . The rectangular preferences of Figure 1a will always choose e , as long as e is Pareto efficient.

Consider, now, the social situation S of Figure 3b. Here, e is not Pareto efficient: both individuals have higher welfare at point f above the 45° line on S 's boundary. In this case, egalitarian preferences choose state f , the state that maximizes the minimum welfare of the two individuals, in the set S . So a more precise name for the rectangular indifference map is 'maximin' preferences. They are the preferences associated with the Rawlsian difference principle, although Rawls did not employ them in a welfarist setting.

Figures 3a and 3b

A number of philosophers, notably Temkin(1993), distinguish between *radical egalitarianism*, which always chooses the point e , Pareto efficient or not, and maximin. I will not consider radical egalitarianism in this paper (because, within a welfarist framework, I find it indefensible²); I retain the description *egalitarian* for the rectangular preferences of Figure 1a.

A prioritarian is one (after Parfit[1997]) who gives priority to the worse off. I take this to mean two things. First, at any point in \mathbf{R}_+^2 that lies above the 45° line (that is, a state in which A is the worse off), the absolute value of the slope of the indifference curve is greater than one and at every point below the 45° line (that is, a state in which B is the worse off), the absolute value of the slope is less than one. Second, as we move away from the 45° line along an indifference curve and above (below) the 45° line, the slope increases (decreases) in absolute value. The first property means that, in order to increase the welfare level of the worse off person by a given increment, the ethical

² I do not say radical egalitarianism is uninteresting, only that it is uninteresting given the axiom of welfarism. The reason that one might prefer strict equality among persons, with respect to some desirable attribute, rather than some unequal state in which everyone has more of that attribute, is that that attribute does not summarize everything that is important to a person or a society.

observer is always willing to trade-off a greater increment of welfare of the better off person. The second property means that the terms of this trade-off become increasingly favorable to the worse off person, the greater is the inequality between the two (in the sense of distance from the 45° line).

The indifference maps in Figures 1b and 1c are both prioritarian. The difference is that the extreme prioritarian of figure 1b is willing to trade off arbitrarily large amounts of the welfare of the better off person to obtain a given increment in the welfare of the worse off person, as the latter approaches a welfare level of zero. The moderate prioritarian of figure 1c, however, places a bound on the degree of trade-off that is permissible. (Mathematically, this means that as we approach the vertical axis along an indifference curve, the slope of the extreme prioritarian's indifference curve approaches infinity, whereas the slope of the moderate prioritarian's indifference curve is bounded.)

Utilitarian ethics are illustrated in Figure 1d. The effect of this indifference map is always to choose the social state in a given situation which maximizes the sum of the welfare levels of the two individuals. The trade-off between the welfare levels is always one for one.

It is clear from Figure 1 that the egalitarian and utilitarian ethics are limiting cases of prioritarianism: the egalitarian will accept *any* trade-off between welfare levels of the better off and the worse off-- he would trade-off the welfare of the better off person against a small gain for the worse-off person at an infinite rate. The utilitarian, on the other boundary, will never trade-off the better-off person's welfare at a rate greater than one-for-one.

While egalitarianism and utilitarianism are each well-defined distributive ethics (that is, they each are associated with one indifference curve map), prioritarianism is a very large family of ethics. There are many families of indifference curves that are prioritarian.

3. Distributional axioms

In this section, I propose some axioms on ethical preferences that I find attractive, and consider the kinds of restriction they impose on ethics.

A1. Ethical preferences are *continuous*.

To understand this axiom, we must first define the *upper* and *lower contour sets* of a social state. The upper (lower) contour set of state u , under a given preference order or ethic, is the set of all states that are weakly preferred³ (dispreferred) to u . Ethical preferences are *continuous* if upper and lower contour sets of every state are closed sets. This means the following. Let $\{u^j \mid j = 1, 2, \dots\}$ be a sequence of states, each of which is in the upper contour set of u , which converges to a state u^* . Then u^* is in the upper contour set of u . A similar statement must hold for lower contour sets.

Informally, continuity means that there are no ‘jumps’ in the ethical preference order. (There would be a jump, if the limit point u^* above were dispreferred to u .) Continuity is not harmless: it eliminates the ‘leximin’ preference order. It is, however, arguably an attractive property.

³ Weakly preferred means preferred or indifferent.

A2. Ethical preferences are *anonymous*.

This means that ethical preferences take into account only the properties of the social situation, not other (extraneous) characteristics of A and B . Of course, the social situation can reflect the fact that, for example, A is handicapped and B is able; but ‘ableness’ only counts to the extent that it is reflected in welfare levels. The mathematical formulation of anonymity is that the ethical indifference curves are symmetric about the 45° line. Another way of saying this is that, if we rotate an indifference curve about the 45° line as the axis, it turns into itself. In particular, a social state (a,c) is always indifferent to the social state (c,a) .

A3a. Ethical preferences are *monotonic*.

This means that if $u_A^1 \geq u_A^2$ and $u_B^1 \geq u_B^2$ then $(u_A^1, u_B^1) \succeq (u_A^2, u_B^2)$. Giving more welfare to people is never ethically undesirable.

Closely related to monotonicity is:

A3b. (Weak Pareto efficiency) If $a > x$ and $c > y$ then $(a,c) \succ (x,y)$.

Thus, if both persons have higher welfare in one state than in another, then the first state is ethically preferred. Without weak Pareto efficiency, we would not have indifference curves, but perhaps indifference *regions* – large sets of states all of which were mutually

indifferent. A3b guarantees that the sets of indifferent states are indeed curves in the plane.

Weak Pareto efficiency is the axiom that permits us to say that ‘higher’ indifference curves are associated with increasing ethical desirability.

We do not insist on (strong) Pareto efficiency⁴, because that would conflict with continuity, given our other axioms.

A4. Ethical preferences are *weakly prioritarian*.

Weak prioritarianism relaxes the requirement that the trade-off that the ethical observer is willing to make, of the better-off person’s welfare against the worse-off person’s welfare, is *increasing* as we move along an indifference curve away from the 45° line; we replace *increasing* with *non-decreasing*. In particular, utilitarianism is weakly prioritarian.

Weak prioritarianism eliminates preferences in which the better off person is favored.

Next comes the first controversial axiom.

A5*. (triage) There is a welfare level $b'' < b$ such that the following holds. Suppose a social situation T in which either A can live a worthwhile life while B dies ‘quite comfortably’ (to be precise with a welfare level at least b''), or A and B both die. Then the first alternative is chosen. Conversely, if in the first alternative described, the welfare

⁴ Pareto efficiency means if $a \geq x$ and $c \geq y$, with at least one equality strict then $(a,c) \succ (x,y)$.

of B is $c < b''$, and in the second alternative, both welfare levels are greater than or equal to b'' , then the second alternative is chosen.

The situation described in the axiom is called triage because the axiom justifies the practice that goes by that name on the battlefield. When medical resources are scarce, they are devoted not to the worst off, who would die even if treated, but only to those who are less badly off, and can be saved. Actually, A5* requires the battlefield nurse first to administer morphine to those who will anyway die, to bring the quality of their death up to b'' . After that, however, the nurse spends time treating only those who will survive if treated, instead of, alternatively, spending time on the fatally wounded, to extend their lives (i.e., increase their welfare) by a small amount. The last sentence in A5* limits the jurisdiction of triage to cases in which those who are left to die are reasonably comfortable. One might also call A5* the lifeboat axiom: if a lifeboat can only save one person, then straws should be drawn to determine who of A and B will jump overboard. (Alternatively, A and B could live for a few days longer in the boat, but eventually both die.)

A5* is illustrated in Figure 4. At state e , both persons die; at state f , B lives a worthwhile life, and A dies quite comfortably, because her utility is greater than b'' . State f must be preferred to state e . We see this rules out egalitarian preferences, which would choose e . (In particular, e is the Pareto efficient equal-utility point in T .)

Indeed, utilitarianism, moderate prioritarianism or extreme prioritarianism might all choose f over e : only egalitarianism is surely ruled out by A4.

It is because of its inconsistency with egalitarianism that triage is a controversial axiom.

Next, I propose:

A6. (universal decency) Suppose a social situation T where there is the possibility of everyone's flourishing:

$$(\exists(u_A, u_B) \in T)(u_i \geq b', \text{ for } i = A, B).$$

Then a state where both people flourish must be chosen.

A6 rules out utilitarianism. To see this, consult Figure 5. Utilitarianism, when faced with the situation T , chooses point g , rather than some point, like f , on the arc \widehat{CD} , the set of states at which both flourish.

figure 5 here

More interesting, however, is the following. A6, A1, and A3 together imply that the indifference curve through (b', b') is rectangular, as shown in Figure 6.

fig 6 here

To see this, consider the sequence of social states $\{u^j\}$ illustrated by the series of points in Figure 6 that converge to \bar{u} . By A6, (b', b') is preferred to every state u^j . Therefore, by continuity (A1), (b', b') must be preferred or indifferent to \bar{u} . But \bar{u} is preferred or indifferent to (b', b') , by monotonicity (A3a). Therefore \bar{u} must be indifferent to (b', b') . This argument works for every point on the rectangular graph illustrated in Figure 6, and so this graph is the indifference curve containing (b', b') .

Unfortunately, the axioms as they now stand are inconsistent. To see this, consider the indifference curve containing $G=(b,b)$, see Figure 7. The indifference curve through (b,b) cannot pass through the interior of triangle FGJ , for if it did, part of it would have slope less than one in absolute value, which is forbidden by A4, since FGJ lies above the 45° ray through the origin. Therefore the indifference curve through point G passes above the segment GJ . But that means that states in the triangle FGJ are dispreferred to G . This directly contradicts A5.

Prioritarianism, in other words, is inconsistent with triage, as stated in A5*. I suggest that to make the axioms compatible, we weaken the triage axiom to:

A5. (restricted triage) Let $T = \{(c,d), (f,g)\}$ where $\min[c,d,f,g] \geq b''$ and $\max[d,f,g] \leq b$. If c is sufficiently large, then (c,d) is chosen. If, however, $d < b''$, or c is not sufficiently large, then (f,g) is chosen.

In T of A5, only A might live a worthwhile life in the first alternative, and both die in the second alternative. We prefer the first alternative only if A 's life, at welfare level c , is sufficiently good: in particular, for A5 to be consistent with A4, we must have $c > 2b-d$, which guarantees that (c,d) lies above the triangle FGJ of Figure 7.

Let us study the indifference curve through (b'',b'') . The shaded region in figure 7 lies in the lower contour set of (b'',b'') , by A5. On the other hand, points to the right of segment EF all lie in the upper contour set of (b'',b'') , by monotonicity. It follows, by continuity, that all the points on the vertical ray \overline{EF} are indifferent to (b'',b'') . By

anonymity, the horizontal ray beginning at $E=(b'',b'')$ is also part of this indifference curve. Therefore the indifference curve containing (b'',b'') is the rectangular curve HH , of the ‘maximin’ type.

It follows that an acceptable indifference curve through (b,b) is the bold curve shown in Figure 7. I have made this indifference curve ‘utilitarian’ for a segment containing (b,b) , because that is the way to give triage the greatest strength. However, the indifference curve cannot intersect the lower indifference curve HH , so it must eventually become asymptotic to HH or to some vertical line to the right of it: so it is not ‘utilitarian’ throughout.

The axioms permit us to fill in the indifference map in a variety of ways in the upper contour set of (b,b) .

I will not analyze what ethical preferences look like in the upper contour set of (b',b') , because I am not particularly interested in the ethics of distribution in a society all of whose members are flourishing. Distributional ethics are only important when it is possible that some people might not flourish. But because the indifference curve through (b',b') is rectangular, all indifference curves above (b',b') lie entirely in the region where both individuals flourish, and so the problem of ethics, there, is (in my view) unimportant.

Finally, I discuss sufficientarianism, which I take to be the doctrine advising the ethical observer to ‘maximize the number of people who have enough’ in any situation.

A complete doctrine requires specifying what ‘enough’ is; let me take it, here, to be the property of living a worthwhile life, that is, having a welfare level greater than b .

AS. (Sufficientarianism) The conjunction of A6 and: If T contains a state in which at least one person has welfare level greater than b , then such a state must be chosen.

Consult figure 8. Four regions are labeled with the number of people who live worthwhile lives in them. It follows from AS that points in region 2 are preferred to points in region 1, which are preferred to points in region 0. Continuity implies that the border between regions 0 and 1, the bold line, constitute one indifference curve and the boundary between 1 and 2, drawn with a lighter line, constitutes another indifference curve⁵. We have a contradiction: the point (b,b) cannot belong to two indifference curves, and there is no way of solving this problem without violating continuity. I suggest that the sufficientarian relax continuity only at the point (b,b) , and assign (b,b) to the lower indifference curve. The indifference maps in the three regions can be completed in a variety of ways.

figure 8 here

The indifference curve (in figure 8) on the common border of regions 0 and 1 is an indifference curve from the family of ‘maximax’ preferences: maximize the welfare of

⁵ The precise argument is as follows. pick two states, P and Q on the bold line. Now choose a sequence of states $\{u^j\}$ in the ‘zero’ region converging to P , and a sequence of states $\{v^i\}$ in the ‘one’ region converging to Q . Since every v^i is preferred to every u^j , by AS, it follows by continuity that Q is weakly preferred to P . Now choose a sequence of states $\{w^j\}$ in the ‘one’ region that converges to P , and a

the best-off person. We see, then, that despite what might appear to be a similarity between the ethic of sufficientarianism and the ethic of triage, they are quite different views. For triage, at least in our formulation, is not inconsistent with prioritarianism, while sufficientarianism is. The lower indifference curve in Figure 8 violates A4 in the sharpest possible way: along it, the observer is willing to make extremely large trade-offs of the welfare of the worse-off individual to increase the welfare of the better-off individual by a small amount. And continuity guarantees that a statement like this is also true in the regions near this indifference curve.

The axioms A1 through A6 do not determine a unique ethic. There are many ways of filling in the indifference map that do not violate the axioms.

One might wish to adopt:

A7. Never let anyone die a horrible death, if possible.

A7 rules out weak prioritarianism. It guarantees that the indifference curves associated with very low levels of ethical desirability are asymptotic to the two co-ordinate axes. By continuity, it guarantees that the pair of co-ordinate axes comprise an indifference curve – the one associated with the least desirable states. (In particular, the state (0,0) is no less undesirable than states $(c,0)$ or $(0,c)$ with $c > 0$: it is no worse to let both people die a horrible death than to let one die a horrible death. This cannot be avoided, without a violation of continuity.)

sequence of states $\{x^i\}$ in the ‘zero’ region that converges to Q . The same argument implies, by continuity, that P is weakly preferred to Q . Hence, P must be indifferent to Q .

In figure 9, I present an indifference curve map that is consistent with A1 through A7. The important point is that there are at least three indifference curves of the egalitarian type: the ones containing $(0,0)$, (b'',b'') , and (b',b') . By continuity, in the regions near those curves, preferences are almost egalitarian⁶. The indifference curve containing (b,b) has a section with slope equal to -1 ; this is done in order to make triage an axiom with some force. Therefore, in the region near (b,b) , preferences are very close to being utilitarian.

In the regions between the three egalitarian indifference curves, we have a good deal of freedom in how to fill in the map. There could be other regions of egalitarianism, for instance. More generally, the indifference map can be of the extreme prioritarian variety. The three egalitarian curves demarcate three regions in which the behavior of preferences are independent, because no two indifference curves ever cross. Thus, we can fill in the lower contour set of (b'',b'') in one way, the region between (b'',b'') and (b',b') in a second way, and the upper contour set of (b',b') in a third way.

4. Conclusion

I hope to have shown that there is no inconsistency in an individual's holding a distributive ethic which appears, sometimes, to be egalitarian, sometimes to be only prioritarian, and sometimes in sharp conflict with egalitarianism, in the sense of advocating triage. Indeed, I believe that axioms like A1 through A7 (excluding AS) are quite compelling, and so an attractive ethic *demand*s that the ethical observer appear to display these 'inconsistencies.' My claim is that these 'inconsistencies' are not true

⁶ This statement means: in a region close to those curves, the indifference map of figure 9 and the egalitarian indifference map are almost indistinguishable.

inconsistencies. The precise resolution involves understanding that one's ethical indifference curve map *can* – and I would say *must*—display quite different features in different regions of the space of possible social states.

References

Frankfurt, Harry. 1987. "Equality as a moral ideal," *Ethics* 98, 21-43

Parfit, Derek. 1997. "Equality and priority," *Ratio* 10, 202-221

Temkin, Larry S. 1993. *Inequality*, Oxford University Press

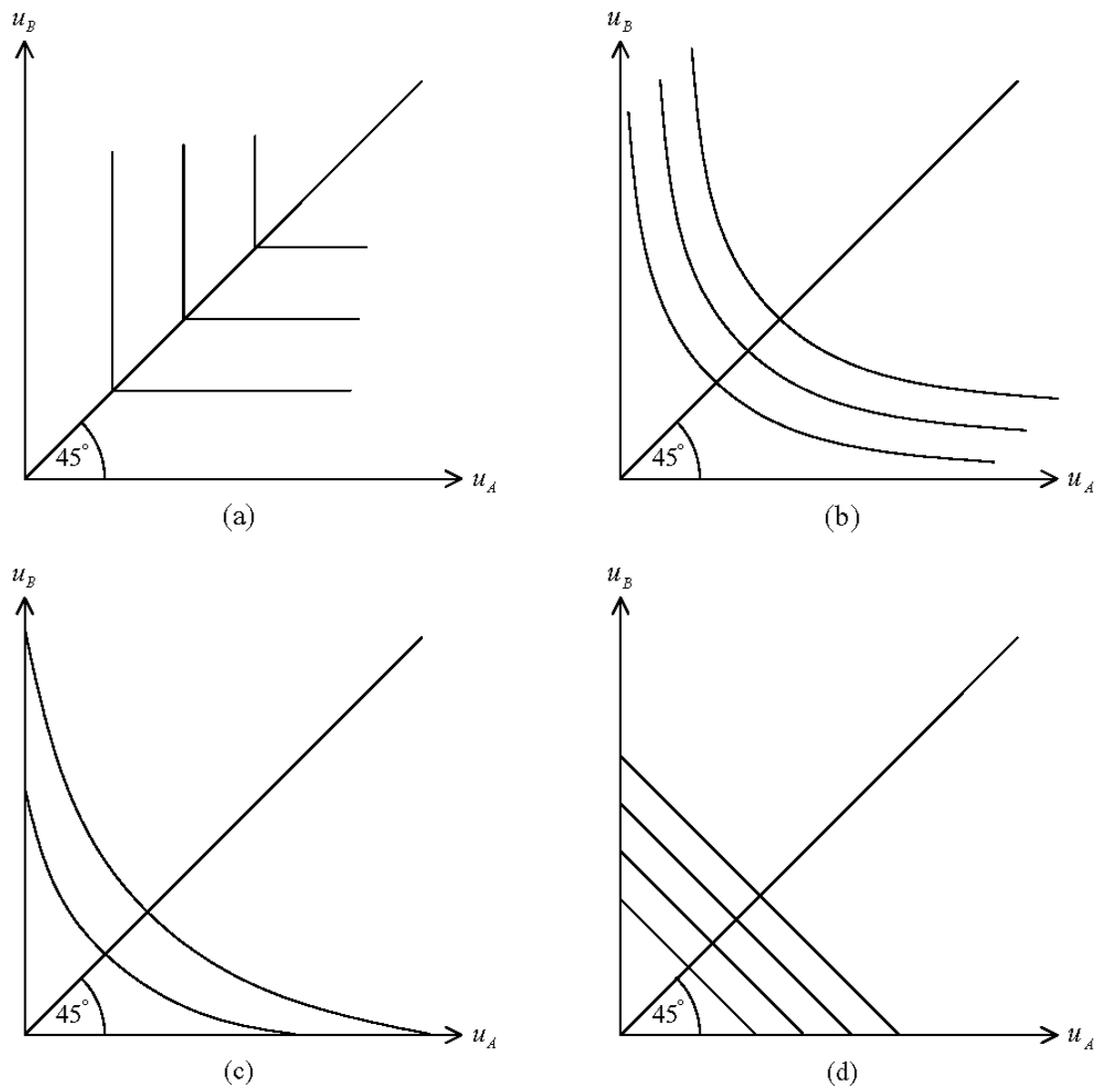


Figure 1 Types of distributive ethic

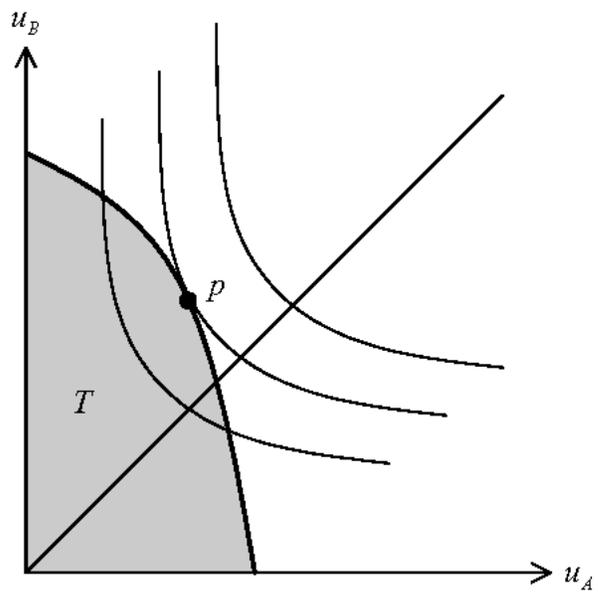


Figure 2 The optimal state

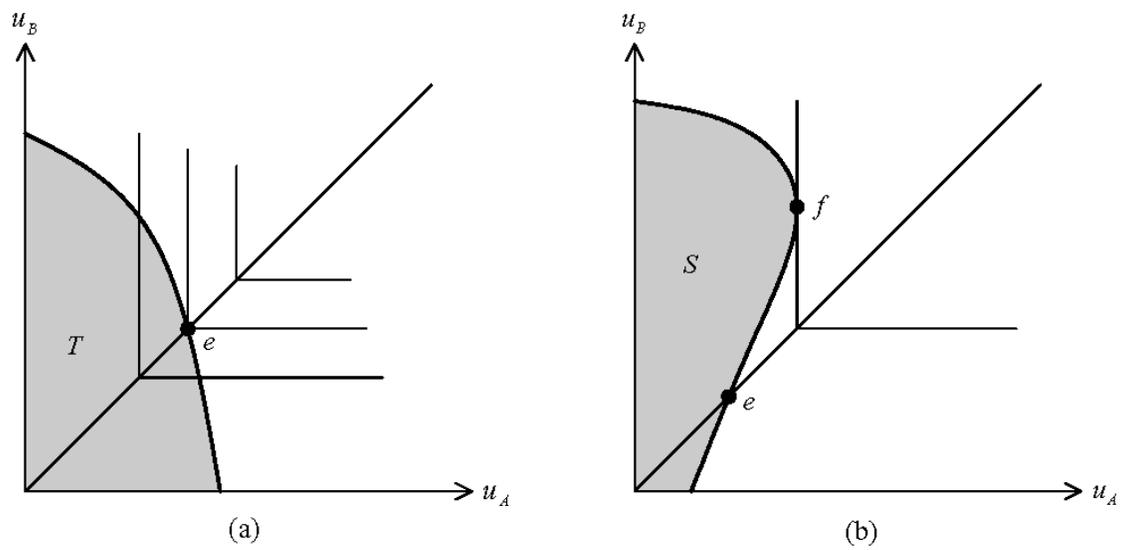


Figure 3 The maximin allocation

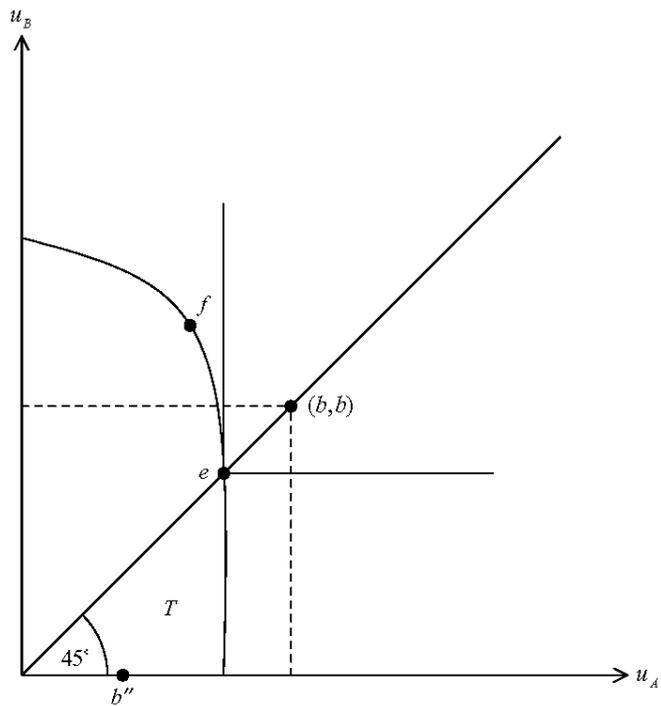


Figure 4 Axiom A5*

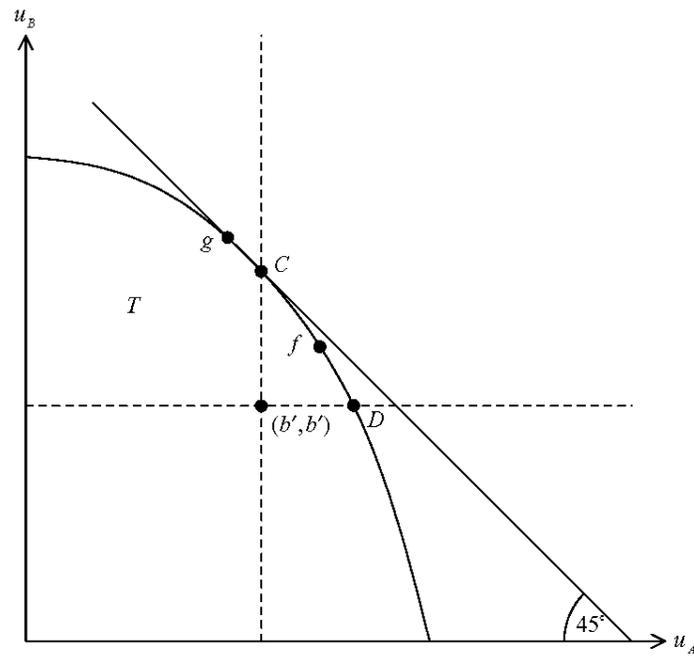


Figure 5 Utilitarianism violates axiom A6

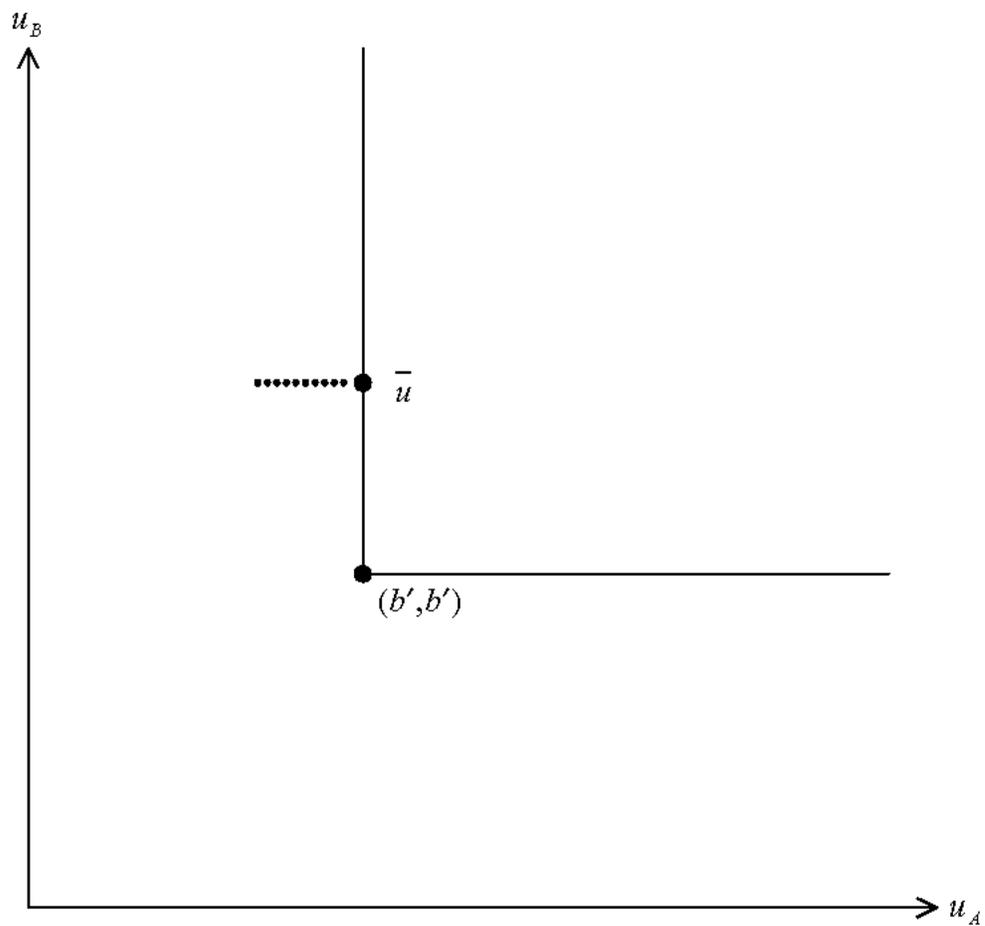


Figure 6 An implication of A1, A3, and A6

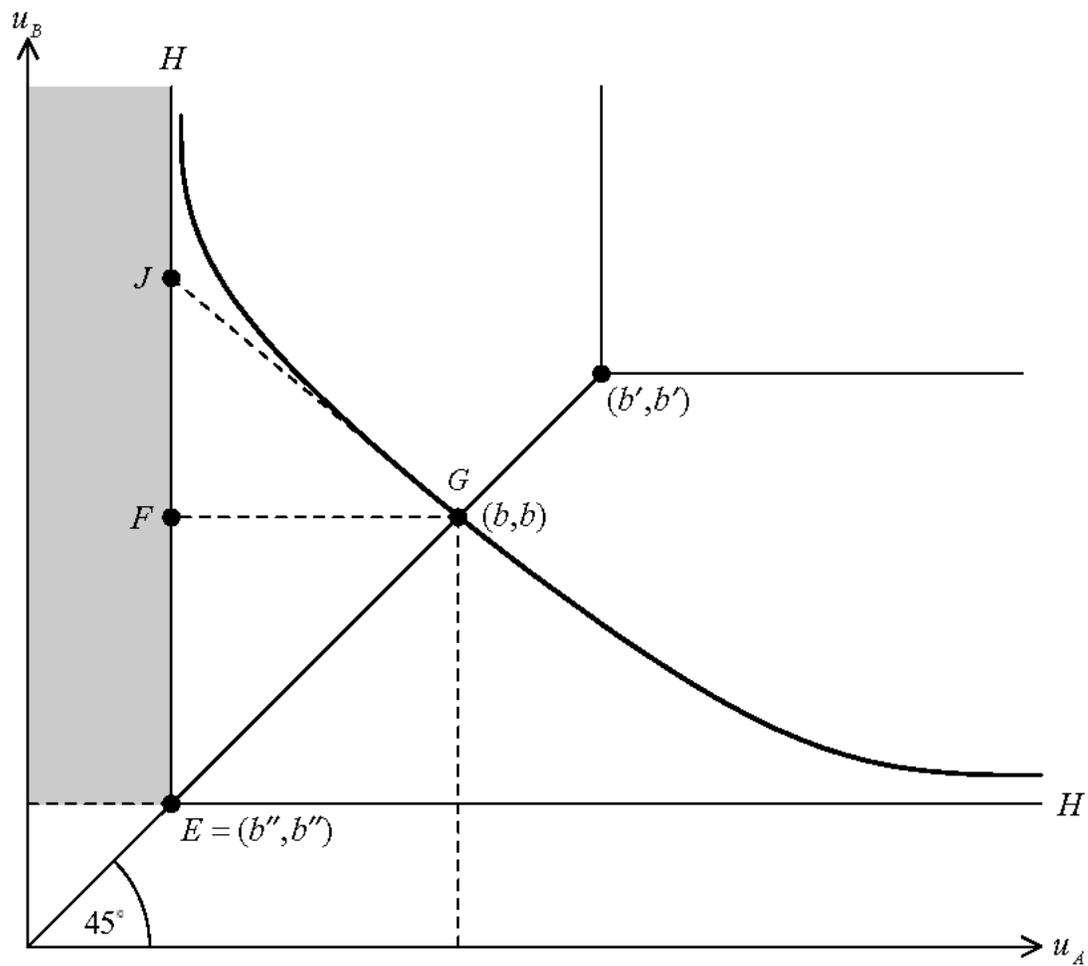


Figure 7 Illustration of an inconsistency

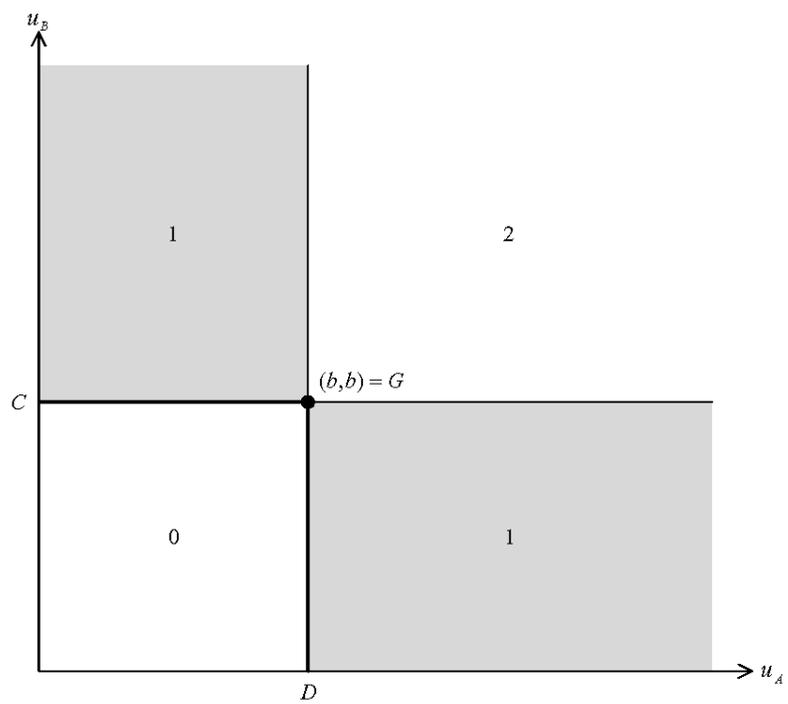


Figure 8 A violation of continuity

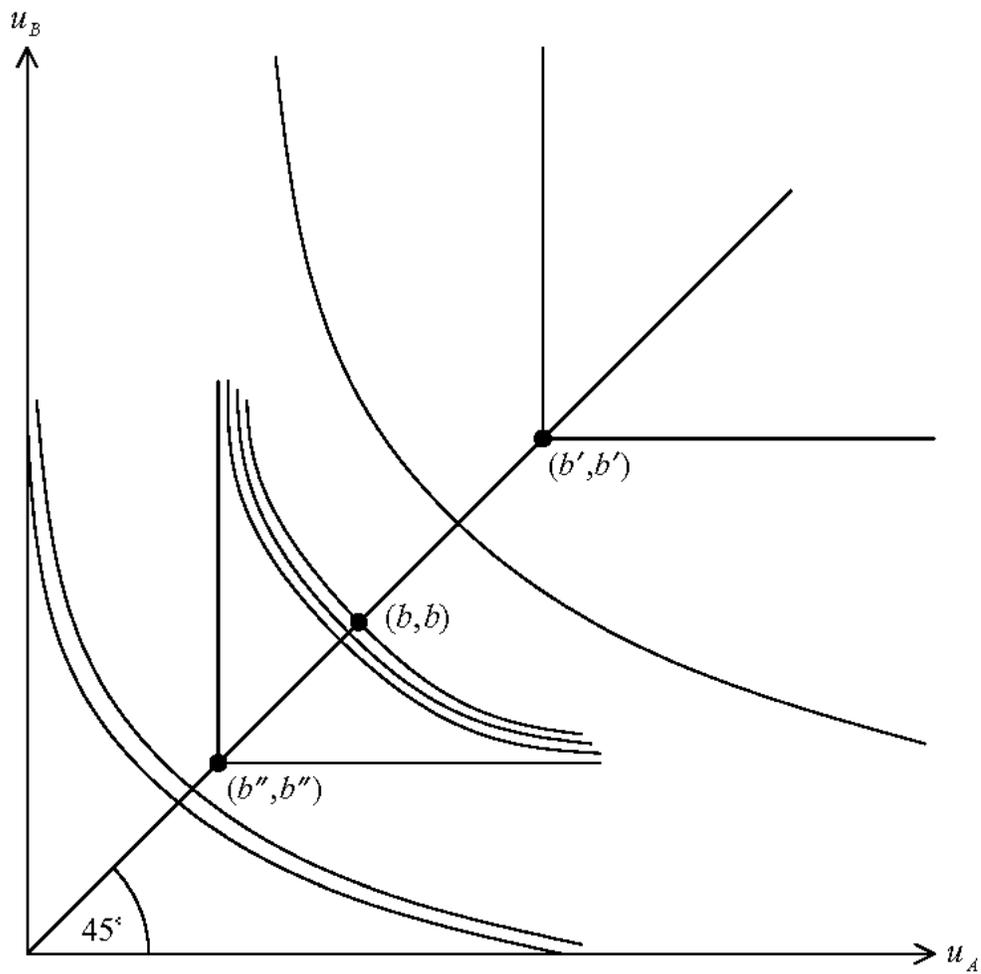


Figure 9 A distributional ethic obeying the axioms