
Justice, Utility, and the Difference Principle

Alan Lundgard

The University of Michigan, Ann Arbor
Ann Arbor, MI
arlu@umich.edu

December 14, 2015

ABSTRACT

Rawls argues that utilitarianism does not take seriously the distinction between persons, whereas justice as fairness does.¹ In some detail, this short paper looks at Rawls's argument for the difference principle. In §1, I reconstruct the indifference curves corresponding to Rawls's well-known criticism that utilitarianism allows the lesser losses of some to be justified by the greater gains of others. In §2, I claim that, insofar as he does not consider diminishing marginal utility, Rawls incorrectly interprets the utilitarian indifference curves. In §3, I show how these curves ought to be redrawn in consideration of diminishing marginal utility, and argue that, once redrawn, the utilitarian picture resembles the difference principle in contexts relevant to distributive justice. Finally, in §4, I conclude that, in light of this resemblance, Rawls's criticism of utilitarianism, although not entirely refuted, might be taken somewhat less seriously.

Keywords Justice, Fairness, John Rawls, Difference Principle, Utilitarianism

1 Rawls's Well-Known Criticism

“The question is whether the imposition of disadvantages on a few can be outweighed by a greater sum of advantages enjoyed by others; or whether the weight of justice requires an equal liberty for all and permits only those economic and social inequalities which are to each person's interests.”²

To better understand the meaning of this quote, consider a simple example. Assume a two-class society and a fixed stock of commodities to be distributed between representative person X_1 and representative person X_2 . Assume further, as Rawls does, that X_1 represents the most advantaged class and that X_2 represents the least advantaged class. Then, says Rawls, “No matter how much either person's situation is improved, there is no gain from the standpoint of the difference principle unless the other gains also.”³ To foreshadow, I will argue in §3 that, in the limit, the utilitarian picture also satisfies the same constraint imposed by the difference principle, but first I'll illustrate just what Rawls means by this constraint.

Observe Figure 1 (The Difference Principle). Suppose there is a fixed stock of commodities to be distributed between two representative persons X_1 and X_2 . Suppose the current distribution is d_1 , corresponding to the indifference curve J_1 , since d_1 lies on that curve. I label the indifference curves J_i , since each point on curve J_i represents a distribution that is “judged equally just”.⁴ Then, the goal is to reach indifference curves further from the origin O . In the language of basic microeconomics, X_1 and X_2 are perfect complements, meaning that increasing X_1 's share of the stock will result in a more just society only if X_2 's share is also increased. Thus, one can never get to J_2 by moving from d_1

¹Rawls, 27

²Ibid., 33

³Ibid., 76: This is not strictly true, since gains could be had via redistribution of X_1 's wealth, but I think Rawls means to say “no matter how much X_1 's situation is improved, there is no gain from the standpoint of the difference principle unless X_2 gains also.”

⁴Ibid., 76

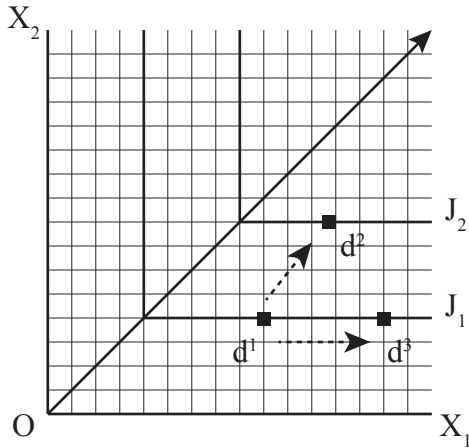


Figure 1: The Difference Principle. In this picture, it is not possible to move from lower to higher indifference curves (from d_1 to d_2) without benefiting the least well-off representative person. Thus, the lesser losses of some are not justified by the greater gains of others.

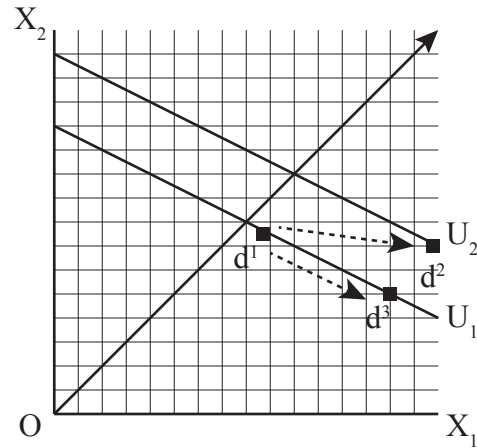


Figure 2: Classical Utilitarianism. In this picture, it is possible that X_2 receives a small loss in moving from d_1 to d_2 , but, by the same token, society overall benefited in moving from U_1 to U_2 . Thus, the lesser losses of some may be justified by the greater gains of others.

horizontally along J_1 to d_3 . (Nor will moving vertically along J_1 allow one to reach J_2). With additional stock (or redistribution of X_1 's share to X_2), the only way to get to J_2 from d_1 is by moving upwards and away from the origin.⁵

Observe Figure 2 (Classical Utilitarianism). Again, suppose there is a fixed stock of commodities to be distributed between two representative persons X_1 and X_2 . Suppose the current distribution is d_1 , corresponding to the indifference curve U_1 , since d_1 lies on that curve. I label the indifference curves U_i , since each point on curve U_i represents a distribution with equal net utility. Similarly, the goal is to maximize net utility by reaching indifference curves further from the origin O . The ideal legislator's goal is to move society from U_1 to U_2 , thus increasing net utility. In the language of basic microeconomics, X_1 and X_2 are perfect substitutes, meaning that the ideal legislator is indifferent to the relative quantities distributed to X_1 and X_2 . Moving in any direction along U_1 (e.g. from d_1 to d_3) will only adjust the relative quantities distributed between X_1 and X_2 . It will not change net utility. With additional stock (or redistribution of X_1 's share to X_2), there will be many ways to get from d_1 to U_2 . Some of these ways could involve losses for X_2 (e.g. in moving from d_1 to d_2 , X_2 is slightly worse off), and, given the linear slope of the indifference curves, this will always be a possibility. Thus, we might expect that Rawls's criticism always holds. Under the obligation to maximize societal utility, the ideal legislator must do so, even if doing so means exchanging lesser losses of some for the greater gains of society. In this way, it is argued, utilitarianism permits violations—however small—of individuals' utilities, and does not take seriously the distinction between persons.

2 The Utilitarian Indifference Curve

Rawls is correct to draw the utilitarian's indifference curves with slopes less than -1 (specifically, I draw them with a slope of $-1/2$).⁶ However, I claim that his reasons for making the curve shallowly sloping are incorrect (to the extent that they lead to unintuitive conclusions). He reasons as follows. Since it can be assumed that X_1 is representative of a population smaller than the population for which X_2 is representative (i.e. since X_1 represents the advantaged class, X_2 represents disadvantaged class, and surely the advantaged class is smaller than the disadvantaged class), a fixed quantity of stock distributed over all persons will result in greater benefit for X_1 than for X_2 . That is, suppose X_2 represents twice as many persons as does X_1 . Make it even simpler and suppose that X_2 represents a class of population size 2

⁵Note: The dashed arrows represent the same possibilities expressed in Rawls's contribution curve (Rawls 76), except that, here, I show them as discrete trade-offs between persons X_1 and X_2 , instead of showing them as a continuous curve, as Rawls does. The idea is the same.

⁶He notes that, if X_1 and X_2 benefited equally from equal quantity of stock, then the indifference curve slopes would be -1 , and each would be perpendicular to the 45-degree line extending from the origin (Rawls 77).

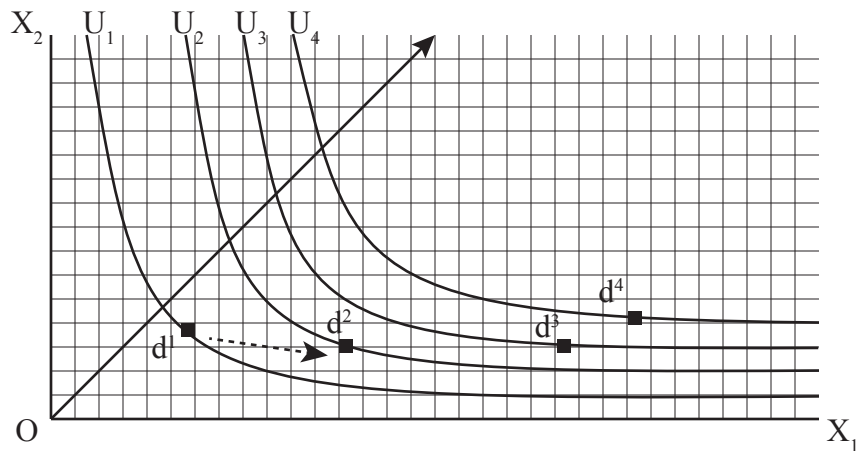


Figure 3: Diminishing Marginal Utility. In this picture, increases to net utility eventually necessitate gains to the least advantaged class X_2 (e.g., in moving from d_3 to d_4), even if small losses are initially permitted (e.g., in moving from d_1 to d_2). Thus, I argue that, in the limit, this picture also satisfies the same constraint imposed by the difference principle.

and X_1 represents a class of population size 1. Then, if the ideal legislator has \$2 to distribute, giving both dollars to X_1 's class gives it's single member a \$2 benefit. But, distributing the \$2 over X_2 's class gives each of its two members only a \$1 benefit. Thus, the \$2 benefits X_1 more than it does X_2 and the slope of the indifference curve should be $-1/2$ (as shown in Figure 2). But this leads to the unintuitive conclusion that \$2 is worth more to X_1 's class (the advantaged class) than it is worth to X_2 's class (the disadvantaged class). If one takes seriously diminishing marginal utility (as, I think, most utilitarians and economists do), then it may easily be the case that X_1 's class gains less utility from \$2 than does X_2 's class, even though X_2 's class is larger. Therefore, although Rawls is correct to draw the curve with a slope shallower than -1 , I claim that he is incorrect in drawing it linearly. Rather, as diminishing marginal utility requires, the indifference curves should be drawn convex to the origin.

3 Diminishing Marginal Utility

Immediately prior to describing the utilitarian indifference curves, Rawls sketches "a view less egalitarian than the difference principle, and perhaps more plausible at first sight" in which the indifference curves are convex to the origin. This convexity, says Rawls, represents "the fact that as either person gains relative to the other, further benefits to him become less valuable from a social point of view."⁷ Rawls is describing diminishing marginal utility. I certainly cannot claim to know all of Rawls's reasons, but I do not (currently) see any that would prevent him from introducing diminishing marginal utility into the utilitarian picture.⁸ So, in the absence of a good reason not to, I'll introduce it, and proceed to argue that doing so diminishes (somewhat) the seriousness of Rawls's criticism that utilitarianism permits the lesser losses of some, in favor of the greater gains of others.

Observe Figure 3 (Diminishing Marginal Utility). The assumptions are the same as those made in Figure 2. The difference is that the indifference curve is now convex to the origin and diminishes rather quickly (for visual effect more so than argumentative necessity). I've also lengthened the window of the graph in order to show the longer-term behavior of the curves. Suppose the current distribution is d_1 , corresponding to the indifference curve U_1 . As before, the ideal legislator's goal will be to get to a distribution on U_2 . It is not necessary, but possible, that the ideal legislator moves from d_1 to d_2 (representing a smaller loss for society's greater gain). Suppose the ideal legislator makes this move. Now that d_2 is the current distribution, suppose some new stock becomes available to the ideal legislator (via increases in productive efficiency or trade, etc.). How will it be distributed? Since X_1 's utility diminishes rather quickly, continuing to decrease X_2 's share (in a manner similar to the move from d_1 to d_2) will only lead to infinitesimal upticks

⁷Ibid., 77

⁸Elsewhere, Rawls (implicitly) relies on diminishing marginal utility in his conception of the maximin situation.

in net utility that will approach zero in the limit. This is effectively a net increase of zero (in the same sense that 1×10^{-a} is effectively zero for all practical computing purposes, when a is a large number). In such a case, it is not possible for the ideal legislator to increase net utility and at the same time decrease X_2 's share. It is possible to hold X_2 's share constant (this is shown moving from d_2 to d_3). But once at d_3 , where will the ideal legislator move to in order to increase net utility? As with the difference principle, there is nowhere to go but up. The upshot of this line of reasoning is the following. The more stock distributed to X_1 , the most advantaged class, the less likely it is that the ideal legislator will be able to increase net utility via a move that involves losses to X_2 . Eventually, that likelihood converges to zero as X_1 's marginal utility converges to zero. Thus, the ideal legislator will eventually be required to increase X_2 's share in order to yield greater net utility.

4 An Appraisal

The consideration of diminishing marginal utility certainly does not entirely save utilitarianism from Rawls's criticism that it allows for, in some cases, the lesser losses (of the less advantaged) to be justified by the greater gains of the rest (including the most advantaged). But it does mean that the criticism cannot be made in all cases. In which cases, then, will it not apply? It won't apply in cases of great inequality between the most and the least advantaged, i.e., where the marginal utility to the most advantaged is vanishingly small. Rawls's criticism of utilitarianism will apply most convincingly when X_1 and X_2 are roughly equally advantaged. But distributive justice is not as concerned with such cases, so we may wonder whether the situation described in Rawls's criticism should be taken as a serious indictment of utilitarianism in its entirety, or somewhat less strongly, as illustrating an edge case that, given the implausible assumption of constant marginal utility, could occur in principle.⁹

Utilitarianism can, in principle, justify lesser losses of some with the greater gains of others. When applied in this way, utilitarianism can also, in principle, reduce to something that looks a lot like the difference principle. This answers the first part of Rawls's question given at the outset (roughly, "Can lesser losses be justified by greater gains?") but it also does some work toward showing that answering this question in the affirmative is not necessarily fatal to the theory. There is a limit on the extent to which lesser losses can be justified by greater gains, as I have argued. I haven't yet answered to the second part of Rawls's question, namely, what does the weight of a utilitarian justice require? Or—perhaps more appropriate to the utilitarian view—to what sorts of distributions does the sentiment of justice attach itself? And further, would the sentiment of justice attach itself to a distribution that was gotten by reducing the share of the least advantaged class? That answer will not likely be found in indifference curves, and looking elsewhere is more than I'm able to attempt in this short paper. I have more than a strong hunch that most utilitarians would argue that the sentiment of justice would not attach itself to such a distribution.¹⁰ I have, at least hopefully, shown that the utilitarian justice (when it takes diminishing marginal utility into consideration), is somewhat less susceptible to the criticism that it would.

⁹Of course, the implausibility of an assumption does not always diminish the theoretical import (or popularity) of the criticism that depends on it. For example, the implausibility of constant marginal utility does not stop some from making the so-called 'utility monster' criticism of utilitarianism.

¹⁰"Good for good is alone one of the dictates of justice...He who accepts benefits and denies return of them when needed inflicts a real hurt by disappointing one of the most natural and reasonable of expectations, and one which he must at least tacitly have encouraged, otherwise the benefits would seldom have been conferred." (Mill 955)

References

- Rawls, John. *A Theory of Justice*. 5th ed. Cambridge: Harvard University Press, 1973. Print.
- Mill, J.S. *Utilitarianism*. Oxford University Press, 2012. Print.
- Freeman, Samuel. *Original Position*. Stanford University, 27 Feb. 1996. Web.
- Harsanyi, John C. *Morality and the Theory of Rational Behavior*. University of California, Berkeley, 1977. Print.
- Rauscher, Frederick. *Kant's Social and Political Philosophy*. Stanford University, 24 July 2007. Web.
- Slote, Michael. *Justice as a Virtue*. Stanford University, 08 Mar. 2002. Web.
- Wenar, Leif. *John Rawls*. Stanford University, 25 Mar. 2008. Web.
- Wilson, Fred. *John Stuart Mill*. Stanford University, 03 Jan. 2002. Web.