

# Any Non-welfarist Method of Policy Assessment Violates the Pareto Principle

---

Louis Kaplow and Steven Shavell

*Harvard Law School and National Bureau of Economic Research*

The public at large, many policymakers, and a number of economists hold views of social welfare that are non-welfarist. That is, they attach some importance to factors other than the effects of policies on individuals' utilities. We show, however, that any non-welfarist method of policy assessment violates the Pareto principle.

## I. Introduction

Economists usually evaluate social policies with reference to individualistic notions of social welfare, under which assessment of policies depends exclusively on their effects on individuals' utilities. This approach of welfarism (so called because it refers to individuals' levels of welfare) may be contrasted to that of many ordinary citizens and government decisionmakers. They tend to think that individuals' utilities are not the only relevant consideration in determining which policies society ought to adopt; they often regard other principles as important in their own right. For example, they may believe that individuals' rewards should be based on merit or desert, that punishment should fit the crime, and so forth.<sup>1</sup>

We thank Lucian Bebchuk, Howard Chang, Eric Maskin, Ari Zweiman, and a referee for comments and the John M. Olin Center for Law, Economics, and Business at Harvard Law School for financial support. Related work of ours includes Kaplow and Shavell (1999) and Kaplow and Shavell (in press), a broad comparison of welfarism and competing normative frameworks, emphasizing literature in philosophy and in law.

<sup>1</sup> Lest we be misunderstood, we are stating that many individuals accord normative weight to such principles, independent of any instrumental value that the principles might have (such as the effect of imposing punishment that fits the crime on deterrence). The view that certain principles are valued per se—regardless of their consequences for individuals'

In addition, a number of economists favor methods of policy assessment that, in whole or in part, diverge from welfarism. For example, Musgrave (1959, 1990), among others, has suggested that, in the evaluation of tax policy, traditional measures of social welfare should be modified to take horizontal equity into account. More broadly, Sen (1985) argues that social welfare should focus on individuals' basic capabilities rather than on their utilities. Indeed, we suppose that many economists believe that it would be reasonable to include some extra-utility elements, of fairness or of justice, in the social welfare function (in addition to concerns about distribution, which welfarism admits).

Given the widely held view that assessment of social policy should not be confined to effects on individuals' utilities, as well as departures from welfarism that have been proposed by economists, we wish to reconsider non-welfarism in general. We show that, for any non-welfarist method of policy assessment (i.e., for any nonindividualistic social welfare function), there always exist circumstances in which the Pareto principle is violated. In other words, any conceivable notion of social welfare that does not depend solely on individuals' utilities will sometimes require adoption of a policy that makes every person worse off. Thus the tension between concern for individuals' well-being and conceptions of social welfare that give weight to factors apart from their well-being is sharper than may have been apparent; endorsement of any form of non-welfarist policy assessment implies that, in certain circumstances, it is socially desirable to make everyone worse off.

This fundamental conflict between non-welfarist methods of policy assessment and the Pareto principle has not, to our knowledge, been demonstrated previously. In a well-known article, however, Sen (1970) showed that a particular form of non-welfarism—under which interference with a stated domain of individuals' activity is not permitted, even if their activity may affect the utility of others—can lead to violation of the Pareto principle. (The explanation for his point is, in essence, that prohibiting the regulation of individuals' activity, when the activity creates negative externalities, may make everyone worse off.)<sup>2</sup> Our conclu-

---

well-being—is, of course, characteristic of the deontological orientation common among moral philosophers.

<sup>2</sup> It should also be noted that Sen interpreted the conflict that he adduced as raising questions about the underlying appeal of the Pareto principle. We, however, are not led to make such an interpretation. Although Sen's criterion (which he calls "liberalism") seems merely to protect a sphere of individuals' activity, it is in fact tantamount to a prohibition against individuals' voluntarily waiving their "rights" to continue their activity in the sphere (in exchange for some concession from others). This implicit prohibition is the source of the conflict with the Pareto principle. Thus the conflict with the Pareto principle identified by Sen can be seen as rooted in a limitation on individuals' rights (rights to contract), not in the protection of individuals' autonomy. We further mention that protecting a sphere of individuals' activity is commonly justified as a means of preventing governmental abuse of power (as we observe in our conclusion), which is a welfarist

sion, by contrast, is general in that it applies to any form of non-welfarism, and, accordingly, the nature of our demonstration is different from his.

## II. Analysis

Let  $x$  denote a complete description of the world. In particular,  $x$  includes a comprehensive account of each of  $n$  individuals' situations and of anything that might be relevant under any method of evaluating the state of the world. Let  $X$  be the set of all conceivable states of the world.

A social welfare function,  $F$ , is a function from the set of states of the world,  $X$ , to the real line,  $R$ .

An individual  $i$ 's utility function,  $U_i$ , is also a function from  $X$  to  $R$ .

An individualistic social welfare function,  $W$ , is a social welfare function of the form  $W(U_1(x), \dots, U_n(x))$ .

The reader may verify that the following statement is true.

**OBSERVATION.** A social welfare function  $F$  is not individualistic if and only if there exist  $x, x' \in X$  such that  $U_i(x) = U_i(x')$  for all  $i$  and  $F(x) \neq F(x')$ .

We remark that familiar conceptions of fairness are associated with social welfare functions that are not individualistic. Consider, for example, the notion that the punishment should fit the crime. If a social evaluator accords weight to this idea of fairness, he would prefer a state of the world  $x$  in which punishments fit crimes to a state  $x'$  in which punishments do not fit crimes, when other things are equal, and thus when all individuals have the same level of utility in the two states. Because, then,  $F(x) \neq F(x')$  even though  $U_i(x) = U_i(x')$  for all  $i$ , the social welfare function that incorporates this notion of fairness in punishment is not individualistic.<sup>3</sup>

The (weak) Pareto principle is that if, for any states  $x, x' \in X$ , we have  $U_i(x) > U_i(x')$  for all  $i$ , then  $F(x) > F(x')$ .

Let us make two assumptions.

**ASSUMPTION 1.** There exists a good such that, if each person has  $\delta$  more of it, then each person is better off. Specifically, let  $m_i$  be individual  $i$ 's amount of the good. Then if two states,  $x$  and  $x'$ , are identical except

---

argument. In fact, when Sen (1992, pp. 144–46) subsequently defends upholding individual rights even at the expense of welfare, he cites problems such as governmental abuse of power and difficulties of enforcing contracts, appeals that are welfarist in nature.

<sup>3</sup>To be concrete, consider the differentiable social welfare function  $F(U_1(y_1), \dots, U_n(y_n), z)$ , where  $y_i$  is the net income of individual  $i$ ,  $z$  is a measure of the expected number of instances in which monetary punishments will not fit an undesirable act, and  $\partial F/\partial z < 0$ . To see that this  $F$  is not individualistic, one need only consider any two states,  $x$  and  $x'$ , in which all individuals have the same level of utility (i.e., income net of any monetary punishments) but in which  $z$  is different.

that, for all  $i$ ,  $m_i$  in  $x$  is higher by  $\delta > 0$  than  $m_i$  in  $x'$ , then  $U_i(x) > U_i(x')$  for all  $i$ .

ASSUMPTION 2. The function  $F$  is continuous in the  $m_i$ .

Observe that assumption 2 does not imply the stronger assumption that  $F$  is continuous in  $x$  and thus does not rule out a variety of non-individualistic social welfare functions that involve discontinuities. (For example, a social welfare function embodying the principle that promises should be kept might fall discontinuously if a promise is broken.) We suppose only that  $F$  is continuous in some good satisfying assumption 1. (We imagine that the social value of at least one ordinary consumption good is unrelated to the normative appeal of promise-keeping, so that an  $F$  reflecting the appeal of promise-keeping would be continuous in the good even if it is not continuous in whether a promise is kept.)<sup>4</sup>

PROPOSITION. If a social welfare function  $F$  satisfies assumptions 1 and 2 and  $F$  is not an individualistic social welfare function, then  $F$  violates the Pareto principle.

*Proof.* If  $F$  is not an individualistic social welfare function, we know from the observation that there exist  $x, x' \in X$  such that  $U_i(x) = U_i(x')$  for all  $i$  and  $F(x) \neq F(x')$ . Suppose, without loss of generality, that  $F(x) > F(x')$ . Construct  $x''$  from  $x'$  by increasing each  $m_i$  in  $x'$  by a positive amount  $\delta$ . By assumption 2 (continuity), we know that if  $\delta$  is sufficiently small, then  $F(x) > F(x'')$ . By assumption 1, we have  $U_i(x'') > U_i(x')$  for all  $i$ , and because  $U_i(x') = U_i(x)$  for all  $i$ , we know that  $U_i(x'') > U_i(x)$  for all  $i$ . Hence, if the Pareto principle is satisfied,  $F(x'') > F(x)$ . But  $F(x) > F(x'')$ , so  $F$  violates the Pareto principle. Q.E.D.

The plausibility of the proposition is suggested by reflection on what it means for a social welfare function not to be individualistic. Such a social welfare function must ascribe weight to some factor independently of its effect on individuals' utilities. Therefore, a social state that is desirable with respect to the (nonutility) factor will be deemed superior to another state that is identical except that (1) it is inferior with respect to the factor and (2) all individuals are slightly better off.

### III. Conclusion

We have demonstrated that any method of policy assessment that is not purely welfarist violates the Pareto principle. That is, policy evaluation that gives any weight to principles independently of their effect on individuals' utilities will sometimes lead to choices under which everyone is worse off. Therefore, if one adheres to the Pareto principle, one must reject any non-welfarist method of policy assessment. It is irrelevant to this conclusion that Pareto dominance will be rare among actual policy

<sup>4</sup> We note that assumptions 1 and 2 are stronger than necessary to prove our result.

alternatives.<sup>5</sup> Belief in the Pareto principle not only rules out the choice of Pareto-dominated policies but also renders inadmissible a wide range of criteria for assessing policy.<sup>6</sup>

Having stressed that policy assessment must be welfarist if one believes in the Pareto principle, let us comment on several senses in which certain apparently utility-independent principles, notably, commonly held notions of fairness, may nevertheless be relevant when evaluating policy under a welfarist approach. First, individuals may have a taste for adherence to a principle of fairness; that is, their utilities might be higher if a policy embodies some notion of fairness (such as that punishment should fit the crime). In this case, the taste for fairness would be relevant under purely welfarist assessment, just as any other taste would. Second, a notion of fairness might be useful for policymaking if it serves as a good proxy principle for raising utilities when direct assessment of policies cannot be undertaken. (Punishing in proportion to the seriousness of crime may lead to approximately optimal deterrence.) Third, some notions of fairness and justice (such as rights of individuals against the government) might usefully be incorporated in rules in order to constrain the behavior of agents who cannot be trusted to use their discretion to maximize social welfare.<sup>7</sup> Fourth, teaching and inculcating principles of fairness and everyday morality are consistent with maximization of individualistic measures of social welfare, for belief in these principles (such as keeping promises) induces individuals to refrain from behavior (breaking promises) that would harm others.<sup>8</sup>

### References

- Diamond, Peter A. "Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparison of Utility: Comment." *J.P.E.* 75 (October 1967): 765–66.
- Hare, Richard M. *Moral Thinking: Its Levels, Method, and Point*. Oxford: Oxford Univ. Press, 1981.

<sup>5</sup> However unlikely it may be that two randomly selected states could be ranked by the Pareto principle, it can be seen from our proof that the following is nevertheless true: At any point (in the space of utility profiles) at which, under a social welfare function, any importance is attached to a non-welfare-based consideration (i.e., wherever, under the social welfare function, two states with the same utility profiles but distinct nonutility information are assessed differently), it is possible to identify other states such that there is a Pareto conflict.

<sup>6</sup> A similar point, that examining the case of identical individuals can lead one to reject certain normative approaches (even if individuals are not in fact identical), is made by Mirrlees (1982).

<sup>7</sup> The libertarian rights examined in Sen (1970)—mentioned in n. 2—and the concern for "process" noted by Diamond (1967) seem appealing for this reason. Indeed, this reason supplies us with an entirely welfarist justification of libertarian and process rights.

<sup>8</sup> In particular, individuals are led to adhere to the principles by a feeling of virtue if they obey the principles and by a feeling of guilt or a fear of social disapproval if they do not. Such instrumental views of morality are emphasized by, among others, Mill (1863), Sidgwick (1907), and Hare (1981).

- Kaplow, Louis, and Shavell, Steven. "The Conflict between Notions of Fairness and the Pareto Principle." *American Law and Econ. Rev.* 1 (Fall 1999): 63–77.
- . *Fairness versus Welfare*. In press.
- Mill, J. S. *Utilitarianism*. London: Parker, Son and Bourn, 1863.
- Mirrlees, James A. "The Economic Uses of Utilitarianism." In *Utilitarianism and Beyond*, edited by Amartya K. Sen and Bernard Williams. Cambridge: Cambridge Univ. Press, 1982.
- Musgrave, Richard A. *The Theory of Public Finance: Study in Public Economics*. New York: McGraw-Hill, 1959.
- . "Horizontal Equity, Once More." *Nat. Tax J.* 43 (June 1990): 113–22.
- Sen, Amartya K. "The Impossibility of a Paretian Liberal." *J.P.E.* 78 (January/February 1970): 152–57.
- . "Minimal Liberty." *Economica* 59 (May 1992): 139–59.
- . *Commodities and Capabilities*. Amsterdam: North-Holland, 1985.
- Sidgwick, Henry. *The Methods of Ethics*. 7th ed. London: Macmillan, 1907.