

# The Conflict between Notions of Fairness and the Pareto Principle

---

Louis Kaplow and Steven Shavell, *Harvard Law School*

Most legal academics and policy makers believe that notions of fairness should be accorded positive weight in evaluating legal policies. We explain, however, that ascribing importance to any notion of fairness (other than one concerned solely with the distribution of income) will sometimes lead to a conflict with the Pareto principle. That is, to endorse a notion of fairness is to endorse the view that it can be desirable to adopt a legal rule that will reduce the well-being of every person in society.

## 1. Introduction

Our general subject concerns normative analysis of social policy, which in the case of the law involves the choice of legal rules. Under normative economic analysis, legal rules are assessed exclusively with regard to their effects on individuals' well-being. Moreover, because evaluations under normative economic analysis are taken to depend positively on individuals' well-being, the Pareto principle would never be violated: if one legal rule were to make everyone better off than another, the latter would never be chosen.

---

We thank Barbara Fried, Christine Jolls, Al Warren, Ari Zweiman, and the referees for comments and the John M. Olin Center for Law, Economics, and Business at Harvard Law School for financial support. This article is motivated by our larger undertaking, "Fairness versus Welfare Economics in Normative Analysis of Law" (which deals at greater length with many issues that are addressed here only briefly). In working on it, we noticed that many notions of fairness conflict with the Pareto principle. That led us to ask whether all notions of fairness might conflict with the Pareto principle, which we here show to be the case.

Send correspondence to: Louis Kaplow, Harvard Law School, 1575 Massachusetts Ave., Cambridge, MA 02138; Fax (617) 496-4880.

Most legal academics and legal decision makers, however, do not base their normative analysis solely on how legal rules affect individuals' well-being. Instead, they usually accord some weight to various notions of fairness.<sup>1</sup> They may endorse specific theories of fairness. For example, in evaluating tort policy, they may attach importance to the classical principle of corrective justice; in considering contract law, to the sanctity of promises; in setting punishment, to the idea of retributive justice. Or they may invoke general conceptions of fairness or justice without reference to particular theories.

In this article, we demonstrate that granting importance to any notion of fairness entails a conflict with the Pareto principle. More precisely, placing any weight on a notion of fairness implies that, in some situations, one will wish to adopt a legal rule that reduces the well-being of every person in society. Thus, the tension between notions of fairness and concerns for individuals' well-being is sharper than has been appreciated. Indeed, because it is our impression that most analysts who accord importance to notions of fairness would not want to contravene the unanimous preferences of the population, they should find our conclusion troubling.

This article is organized as follows. In section 2, we state the Pareto principle, elaborate on what we mean by notions of fairness, and explain that concerns about the overall distribution of income are not included in the notions of fairness that we consider. In section 3, we provide a general argument showing that any notion of fairness will sometimes conflict with the Pareto principle. In section 4, we illustrate the conflict in several basic areas of law. In section 5, we amplify on the point that the conflict with the Pareto principle constitutes a serious problem for proponents of notions of fairness. (Among other things, we explain why the problem is not attenuated by the fact that, in reality, the choice among legal rules will rarely involve unanimous preferences.) In section 6, we conclude.

---

1. Some legal academics would accord complete weight to notions of fairness in their normative assessment of legal rules. That is, like purely deontological moral philosophers, they believe that rules should be chosen entirely on the basis of principles of fairness, without explicit regard to the effects of the rules on individuals' well-being.

## 2. The Pareto Principle and Notions of Fairness

Because our discussion of the Pareto principle and of notions of fairness makes reference to individuals' well-being, we should state what we mean by that concept. We have in mind the broadest conception of well-being, which includes not merely individuals' levels of material comfort but also their degree of aesthetic fulfillment, their feelings for others, and anything else that they might value, however intangible. For convenience, we assume that individuals understand fully how various situations affect their well-being, so that when we say that an individual is better off, there will be no ambiguity about our meaning.<sup>2</sup>

### Pareto Principle

Under the Pareto principle, if every person is better off under one policy than under another, the former policy is deemed to be socially preferable.<sup>3</sup>

### Fairness

By conceptions of fairness, we refer to principles that are employed in the evaluation of legal rules and that are based, at least in part, on factors unrelated to individuals' well-being.<sup>4</sup> An implication of this definition is that knowledge of the effects of a legal rule on each individual's well-being will not generally be sufficient to assess the fairness of the rule.<sup>5</sup>

---

2. If individuals do not understand how situations affect their well-being, our argument may be applied (without any other modification) to individuals' actual well-being rather than to individuals' well-being as reflected in their mistaken preferences.

3. This principle is formally referred to as the weak Pareto principle. Under the strong version of the principle, if everyone is at least as well off and at least one person is better off under one policy than under another, then the former policy is deemed to be socially preferable. (Our argument in section 3, that any notion of fairness conflicts with the weak Pareto principle, implies a conflict with the strong version as well.)

4. We mean our definition of conceptions of fairness to include principles that are based on factors that matter to individuals but under which the factors are weighed differently from the manner in which individuals themselves weigh them. For example, a principle that says punishment should be proportional to the harm caused by an act depends on factors that matter to people: punishment matters to those punished and, because of deterrence, to others; and harm matters to victims. But under the principle in question, punishment is not assessed solely with regard to how it affects individuals' well-being. (For a formal statement of our definition of a notion of fairness, see note 5.)

5. Our definition of a notion of fairness can be stated formally using the apparatus of welfare economics. First, consider a method of policy assessment that depends only

For example, under the classical conception of corrective justice, it is important in itself that an individual who wrongfully injures another be required to compensate him.<sup>6</sup> That is, the compensation requirement is treated as important, apart from any effects it may have on individuals' well-being, notably, that the prospect of having to pay compensation may deter wrongful behavior and that the availability of compensation may serve as implicit insurance for risk-averse victims. Likewise, under the view that keeping promises is valuable in its own right, contract rules should sometimes prevent breach even when breach would otherwise raise individuals' welfare.<sup>7</sup> Or, under the principle of retributive justice, weight would be given to the goal that punishment be in proportion to the gravity of wrongful acts, independently of the effect of such punishment on the level of wrongdoing and of the costs of its imposition.<sup>8</sup>

We note that, in defining notions of fairness as normative principles that are not based exclusively on individuals' welfare, we are hardly suggesting that giving weight to most notions of fairness will tend to favor policies

---

on individuals' well-being. Specifically, suppose that there are  $n$  individuals, and let the well-being (utility) of the first individual be denoted  $U_1$ , that of the second  $U_2$ , and so forth. Also, let  $x$  stand for an exhaustive description of a situation, interpreted as a state of the world that will prevail under some legal regime. Then social welfare,  $W(x)$ , can be written as

$$W(x) = F(U_1(x), U_2(x), \dots, U_n(x)),$$

where  $W(x_1) > W(x_2)$  is taken to mean that situation  $x_1$  is socially preferred to situation  $x_2$ . We note that economists often call  $W$  an individualistic social welfare function because social welfare depends only on each individual's well-being. We also note that  $W$  is ordinarily taken to be an increasing function of the  $U_i$ . A method of policy assessment that gives weight to a notion of fairness corresponds to a function  $Z(x)$  that differs from (cannot be expressed in the form of) the  $W$  function just defined. Consider

$$Z(x) = F(U_1(x), U_2(x), \dots, U_n(x), x).$$

Here,  $Z$  may depend not only on each individual's utility, but also directly on  $x$ , which includes all characteristics of the situation that will prevail under a legal regime. Thus, it is possible that a characteristic of the situation  $x$  that affects no one's utility nevertheless affects  $Z$ . Moreover, it may be that a characteristic of a situation  $x$  that affects individuals' utilities influences  $Z$  but in a different manner. For further discussion of the formal difference between  $Z$  functions and  $W$  functions, see Kaplow and Shavell (1999).

6. See, for example, Coleman (1992) for a modern discussion of the Aristotelian notion of corrective justice.

7. See, for example, Ross (1930) on promise keeping.

8. For a survey, see Benn (1967).

that systematically reduce individuals' well-being. It is obvious that many notions of fairness will frequently favor policies that enhance individuals' well-being (for example, a requirement that injurers compensate victims may often beneficially create deterrence). Moreover, those who accord weight to notions of fairness in evaluating legal rules typically do not view fairness as absolute; instead, they often hold mixed views, under which weight is given both to notions of fairness and to the effects of legal rules on individuals' welfare (if the administrative cost of a rule requiring compensation were exorbitant, they might not insist upon it). Nevertheless, because notions of fairness depend on factors that are independent of individuals' well-being, according weight to fairness will sometimes lead one to choose different legal rules from those that would be selected if rules were chosen with sole regard to individuals' well-being. These instances, in which different rules would be chosen, are the focus of our discussion in the sections to follow.

### Remark on Concerns about the Distribution of Income

Because our argument that notions of fairness conflict with the Pareto principle will be seen to depend critically on the assumption that fairness notions are not based exclusively on individuals' well-being, our argument does not apply to many conceptions about equality in the distribution of income. The reason is that some such concerns about the distribution of income are incorporated in normative analysis that is based solely on individuals' well-being.<sup>9</sup> For example, the familiar idea that income should be redistributed from the rich to the poor, because the marginal utility of income is greater for the poor, is included in normative analysis based only on individuals' well-being.<sup>10</sup>

---

9. There also exist theories bearing on the just distribution of income that are not based solely on individuals' well-being, and these theories can be shown sometimes to favor policy choices that conflict with the Pareto principle. (We do not demonstrate this below because our analysis is confined to cases in which income distribution is not at issue. But a more general proof of our proposition about the conflict between the Pareto principle and notions of fairness, in Kaplow and Shavell [1999], encompasses distributive theories that are not based only on individuals' utilities.)

10. To see how the distribution of income may be relevant to normative assessment that is based only on individuals' well-being, let us consider examples of individualistic social welfare functions,  $W$ , defined in note 5. If  $W$  were the utilitarian social welfare function (the sum of individuals' utilities) or a similar aggregative function, then redis-

### 3. The Conflict between Notions of Fairness and the Pareto Principle

To demonstrate that there is a conflict between notions of fairness and the Pareto principle, we will consider contexts in which all individuals are symmetrically situated before events arise that call for the application of legal rules.<sup>11</sup> Thus, with regard to automobile accidents, we envision that everyone is equally likely to be an injurer or a victim; with respect to contracts, that each individual is just as likely to be a breaching party or a victim of breach; and so forth. (In section 5, we will elaborate on why such contexts are particularly important to consider.)

Observe first that, in the contexts to be examined, everyone will favor the same legal rule because everyone begins in the same situation.<sup>12</sup> In other words, using the Pareto principle, one can determine which legal rule is best among any set of legal rules that might be contemplated.

We now explain why it follows, almost automatically from the definition of a notion of fairness, that assessment of legal rules based on a notion of fairness will sometimes result in the choice of a legal rule that makes everyone worse off.<sup>13</sup> Consider any notion of fairness and observe that, as we mentioned in section 2, it will in some instances lead to the selection of legal rules different from those that would be chosen if one were exclusively concerned with individuals' well-being.<sup>14</sup> But, in the cases that

---

tributing income from the rich to the poor, as described in the text, would raise social welfare. In addition, if  $W$  were to depend directly upon the distribution of individuals' utilities, then the distribution of income would be important, apart from how it affects the sum of utilities. For example,  $W$  might weigh increases in the utility of worse-off individuals more heavily than decreases in the utility of better-off individuals.

11. More precisely, we assume that everyone has the same income and preferences, is exposed to the same risks, and has the same opportunities.

12. The statement in the text reflects a prospective view of what individuals would favor, before they learn whether they in particular will happen to gain or lose under a particular legal rule. Although we believe that this prospective view of individuals' preferences is compelling, we note that our argument can also be made by considering only contexts in which individuals' preferences are also unanimous *ex post*. For example, we could suppose that each person will definitely be once in the position of injurer and once in that of victim. In such a context, the *ex ante* and *ex post* perspectives are identical: each individual's expected utility equals the actual level of utility he will ultimately experience.

13. For a formal and more general proof, see Kaplow and Shavell (1999).

14. To amplify, note that if a notion of fairness were always to favor the legal rule that would be chosen if one were exclusively concerned with individuals' well-

we are examining, to select legal rules with exclusive regard for individuals' well-being is to choose the rule that makes all individuals best off. Therefore, whenever the notion of fairness would, as it sometimes must, lead to the choice of a different rule, all individuals will be made worse off. That is, the notion of fairness will conflict with the Pareto principle in some situations.

The conflict between notions of fairness and the Pareto principle can also be shown in a different manner. Suppose that, under a notion of fairness, legal rule *A* is deemed preferable to legal rule *B* because rule *A* is better in some respect that is independent of individuals' well-being. If everyone happens to be worse off under rule *A*, then the notion of fairness conflicts with the Pareto principle and we are done with our argument. If, instead, everyone is better off under rule *A*, we can show that everyone will be worse off under a modified rule *A* that is still viewed as more fair than rule *B*. In particular, consider a modified rule *A* that is identical to the original rule *A* except that its administrative costs are higher by enough that everyone is worse off under modified rule *A* than under rule *B*. However, this modified rule *A* is in all respects other than administrative costs identical to the original rule *A*, which, in turn, was stated to be more fair than rule *B* on some ground unrelated to individuals' well-being; therefore this modified rule *A* also must be more fair than rule *B*, even though the latter makes everyone better off.<sup>15</sup> Thus, we

---

being, it must do so in all possible regimes—whatever are individuals' preferences, the available set of legal rules, and the other characteristics of the symmetric situation (such as the set of activities that are available, the costs of undertaking them, and the harm they impose on others). The only way a notion of fairness could do this would be if its import varied in such a manner that the rule chosen under fairness-based analysis changed at precisely the point that individuals' welfare assessments changed in a way that led them to favor a different rule, and this would need to be true across all possible regimes. However, because notions of fairness by definition attach importance to factors independent of individuals' well-being, that is not possible. See also notes 15 and 16.

15. We emphasize that the argument in the text is applicable to a notion of fairness (perhaps a mixed notion) under which the evaluation of legal rules may also be affected by administrative costs. To see this, consider the situation in which administrative costs under modified rule *A* are just at the level at which individuals are indifferent between it and rule *B*. In this case, modified rule *A* must be preferred to rule *B* under the fairness criterion because, by assumption, rule *A* was better than rule *B* on account of a factor apart from individuals' well-being. Hence, if administrative costs under modified rule

have constructed a case in which the notion of fairness conflicts with the Pareto principle.

#### 4. Illustrations

##### Corrective Justice and Tort Law

In this illustration and those to follow, we continue to assume that individuals are symmetrically situated at the outset. Let us consider automobile accidents, supposing that each driver is equally likely to be an injurer or a victim, and compare strict liability and the negligence rule. Assume that the negligence rule is deemed to be more fair than strict liability because corrective justice demands that liability be limited to wrongdoers, understood to mean individuals who were at fault. One possibility is that everyone is better off under strict liability in the situation under examination. If this is the case, allowing corrective justice to guide legal policy, which entails use of the negligence rule, would make everyone worse off. The other possibility is that in which everyone is better off under the negligence rule than under strict liability. Following the argument in section 3, we can nevertheless construct a situation in which adhering to corrective justice would conflict with the Pareto principle. Namely, consider a modified situation in which the negligence rule involves greater administrative costs. If these administrative costs are sufficiently higher, then the modified negligence rule, despite being more just, will make everyone worse off than they would be under strict liability.<sup>16</sup>

---

*A* were raised slightly more, the fairness criterion would still favor modified rule *A* even though rule *B* makes everyone better off. (For a numerical example, see note 16.)

16. As we explained in note 15, the argument in the text is applicable to a mixed notion of fairness that accords weight to factors, such as administrative costs, that affect individuals' well-being. To elaborate, suppose initially that the level of individuals' well-being under the negligence rule is 100, that the level under strict liability is 90, and that the weight given by an analyst to fairness (in favor of the negligence rule) is 5. Thus, the advantage of the negligence rule in terms of individuals' well-being is 10 and that in terms of fairness is 5, so the fairness-oriented analyst would prefer the negligence rule. Now, as the administrative costs of the negligence rule are increased, there will come a point after which well-being under the negligence will be below that under strict liability. Suppose that the administrative costs are raised to a level at which individuals' well-being under the negligence rule is 88. Then everyone is better off under strict liability—individuals' well-being is higher by  $90 - 88$ , or 2—but the fairness-oriented analyst would still choose the negligence rule because fairness is



## Fair Punishment and Law Enforcement

Next consider littering and suppose that, under a notion of retributive justice, the fitting punishment is a \$50 fine. Suppose, however, that if the fine were raised to \$75, it would be possible to reduce the number of officers required to police littering while maintaining the same level of deterrence (litterers would be caught less often, but this would be offset by their having to pay the higher fine when caught). Assume that using the \$75 fine results in a net improvement in everyone's well-being because everyone's taxes are lower due to the reduced need for policing and that this benefit exceeds any costs that may arise under a regime with higher fines.<sup>17</sup> Thus, under the fair, \$50 level of punishment, everyone expects to be worse off than under the higher, \$75 level of punishment.<sup>18</sup>

## Right to Legal Redress and Litigation

Here we examine situations in which individuals are equally likely to be plaintiffs and defendants, and we consider a notion of fairness under which any plaintiff with a valid claim has a right to legal redress. It is apparent that, in some settings, everyone will be better off if this right is not upheld. On one hand, individuals will benefit under a regime that upholds the right because the prospect of suit may create deterrence and it may provide implicit insurance for risk-averse victims. On the other hand, individuals will incur costs, notably, litigation expenses. When these costs are sufficiently high, they will outweigh the benefits of suit, in which case everyone will be better off if the right to sue is not upheld.<sup>19</sup>

---

given a weight of 5, which exceeds the benefit of strict liability in terms of individuals' well-being, 2. (We note that this argument holds even if the weight given to fairness is very small. If, say, the weight is only 1, administrative costs under the negligence rule could be raised somewhat less than in the prior case, to the point at which well-being under the negligence rule is 89.5.)

17. Such costs might stem from more frequent challenges to citations for littering and greater imposition of risk.

18. As in our preceding example, see *supra* note 16, an analyst who holds a mixed view would give weight to individuals' improved well-being from the higher, unfair fine; but, if fairness is given any weight, there will exist situations in which the analyst will continue to favor setting the fine at the fair level even though everyone is worse off than under the higher-fine alternative.

19. Also, as we have noted in our prior examples, see *supra* note 16, even if an analyst gives weight both to fairness and to individuals' well-being, there will exist situations in which he will choose to uphold the right to redress even though all individuals will be made worse off.

## 5. Implications and Importance

### Logical Consistency

We have shown that any notion of fairness will conflict with the Pareto principle in certain situations. Thus, as a matter of logical consistency, if one is to adhere to the Pareto principle, one cannot give any weight in normative analysis to notions of fairness. If a theory, in some part of its domain of application, contradicts a principle to which one subscribes, the theory has thereby been rejected.<sup>20</sup>

Our conclusion should be related to the point that, in reality, it will rarely be the case that one legal rule will literally make everyone better off than will another legal rule. This observation has led most analysts to believe that the Pareto principle has little relevance in making policy decisions. What our argument shows, however, is that even though adherence to the Pareto principle may not directly determine policy *choices* in most real situations, it nevertheless has powerful implications for what *criteria* for making policy choices one can consistently employ.

### Some Important Sources of Appeal of Notions of Fairness

First, it seems that concerns about aiding or protecting individuals underlie many analysts' beliefs in notions of fairness.<sup>21</sup> Under corrective

---

20. It has been suggested to us that perhaps the conflict between a notion of fairness and the Pareto principle could be avoided if the notion were modified by assuming it to be inapplicable whenever it would conflict with the Pareto principle. (We note that such a modified notion of fairness is not advanced in any legal or philosophical literature of which we are aware and, in any event, would be inconsistent with the spirit of the rationales offered for notions of fairness.) Such a modified notion of fairness would, however, be odd because, under it, trivial changes in facts would alter the entire basis for assessing legal policies. To illustrate, suppose that in regime *Y* everyone is better off than in regime *Z*, but that one person is only very slightly better off, say, by a penny. Now consider a regime *Y'* in which the only difference from *Y* is that this person is not better off but instead is worse off than in *Z* by a penny. Then, under the modified notion of fairness, the normative method of evaluation would change abruptly—from the Pareto principle, based solely on individuals' well-being, to some notion of fairness—as a consequence of a two-cent difference in one person's position. (Note that, under the modified notion of fairness, regime *Z* is assessed differently depending on which regime it is being compared to—in this case, *Y* or *Y'*.) See also the formal treatment of fairness and continuity in Kaplow and Shavell (1999).

21. We recognize that some individuals, notably, purely deontological moral philosophers, do not claim to find notions of fairness appealing on this sort of ground.

justice, for example, the requirement of compensation for wrongful injury aids victims, and the restriction of compensation to cases of wrongdoing protects injurers who did not act inappropriately. Under retributive justice, the limitation of sanctions to levels that are in proportion to the gravity of acts protects violators against excessive punishment. And so forth.

Our demonstration that all notions of fairness will conflict with the Pareto principle in certain situations means, however, that it cannot be an inherent feature of a notion of fairness that some individuals are helped. When there is a conflict with the Pareto principle, everyone is worse off, including any person of possible concern under a notion of fairness.

Second, many notions of fairness are grounded in ideals of personal freedom and autonomy or libertarian conceptions of individual rights and liberty. But our demonstration implies that consistently adhering to such notions will sometimes require adopting a policy that would be rejected by every individual if in fact each was permitted freely to make his own choice about the policy. If, however, all individuals would want to reject the policy, it would seem that adopting the policy would interfere with their freedom or liberty. Thus, there appears to be a problem with the internal logic underlying these notions of fairness.<sup>22</sup>

### Special Relevance of the Symmetric Case

In order to explain the conflict between notions of fairness and the Pareto principle, we examined symmetric settings in which individuals were identically situated at the outset. Here, we emphasize that such settings are especially relevant to important moral theories that are employed to support notions of fairness because these theories make use, explicitly or implicitly, of symmetric settings.

Consider some of the more prominent viewpoints that are advanced when reflecting upon moral principles. The Golden Rule asks that we as-

---

22. This point bears on an argument of Sen (1970). He suggests that adherence to the Pareto principle may be problematic because it conflicts with a certain notion of individual liberty, under which each individual has absolute control over some domain of his activity. (The reason that a Pareto conflict can arise in Sen's construction is that each person's action in his protected domain imposes an externality on others; thus, stipulating that an externality-causing action may not be regulated might make everyone worse off.) But satisfying Sen's libertarian notion of fairness implicitly entails the government's forbidding individuals from voluntarily waiving their rights in exchange for others' waivers, which would be to their mutual benefit.

sess contemplated acts under the hypothetical assumption that others will behave toward us in the same manner. Kant's categorical imperative similarly tests the morality of rules by asking whether they would retain appeal if we generalized them to all situations, particularly including the case in which we ourselves would be in the opposite position from our own, actual one. Harsanyi and Rawls suggest that, in assessing social regimes, we imagine ourselves to be behind a veil of ignorance, not knowing which role we will in fact occupy.

These viewpoints are designed to engender a disinterested outlook in us. This can be accomplished by having us imagine that we are equally likely to be in each possible role or, equivalently, that we will occupy each and every role. To create the proper perspective for normative evaluation, we are asked hypothetically to construct the sort of symmetric situation that has been examined in this article.<sup>23</sup> But, as has been shown here in such situations, when notions of fairness favor legal rules different from those favored under an approach that confines attention to individuals' well-being, there will *always* be a conflict with the Pareto principle. Thus, our argument poses a particularly sharp challenge for analysts who find compelling the sorts of moral theories under consideration.

### Significance of Notions of Fairness Other than as Independent Normative Principles

We have thus far employed the term "fairness" to refer to independent normative principles that are used to assess legal rules. We now consider three senses in which notions of fairness may be of significance other than as independent normative principles.

First, a notion of fairness might be given weight by a legal policy maker because it may correspond to a *taste* that individuals have. For example,

---

23. To understand why principles like the categorical imperative essentially require an assumption of symmetry, consider an example. Suppose that there are strong people and weak people. A strong person would be happy to live by the principle that "might makes right" because this would be to his advantage. In the absence of a symmetry assumption, however, the categorical imperative does not interfere with his adopting this principle because, if it were adopted as a general rule for society, it would still be to his advantage, because he is strong. In order for the categorical imperative to rule out principles like "might makes right," which are little more than statements of self-interest, it is necessary either to imagine that everyone is identical (neither strong nor weak) or that each of us is in each position (strong and weak) for a commonly specified period of time, and this is precisely what is achieved by the symmetry assumption.

people might be upset by the imposition of excessively harsh sanctions, or they might be angry if those who commit heinous crimes are not punished adequately. Evaluation of legal rules that is based on individuals' well-being credits such tastes just as it credits any other tastes. (Note that the existence and significance of tastes for notions of fairness is an entirely empirical matter, not one of moral principle.<sup>24</sup>) It should be observed that the relevance of tastes for fairness when evaluating legal rules does not lead to conflict with the Pareto principle, because that principle depends on individuals' well-being, which itself is now assumed to be directly affected by the tastes.<sup>25</sup>

Second, a notion of fairness might be employed by a legal policy maker as a convenient *proxy* for advancing individuals' welfare. For example, we have mentioned that a requirement that wrongdoers compensate their victims might usefully deter harmful acts. When there is little rigorous analysis or empirical evidence to draw upon, a policy maker may be justified in referring to notions of fairness if it is plausible that they will tend to lead to the choice of policies that make individuals better off (even though, of course, this will not always be the case). Obviously, this is an observation

---

24. By contrast, notions of fairness that are treated as having independent normative significance—the chief concern in this article—affect the evaluation of legal rules regardless of whether individuals have a taste for them. In most philosophical writing and, apparently, in legal writing as well, those who reason using notions of fairness do not regard themselves as merely making empirical conjectures about what individuals' tastes happen to be.

25. In particular, if a legal rule is chosen because of individuals' taste for the rule, it must be that the importance to them of this taste exceeds the cost to them of any of its disadvantages (such as higher administrative costs, less deterrence). Therefore, it must be that individuals are better off under the rule and the Pareto principle is not violated. This case differs from the one we have been emphasizing throughout the article, where a notion of fairness is given weight as an independent normative principle and thus may affect the choice of rules regardless of whether individuals have a taste for the notion. We now mention a separate point. Our discussion of fairness as a taste raises the question whether, by simply *imputing* to all individuals a common, "correct" taste for a notion of fairness, one could avoid the conflict we have identified between the Pareto principle and giving weight to notions of fairness as independent normative principles. Under this view, all might be deemed to have a taste for a notion of fairness such that, given their deemed well-being, they all would rank legal rules identically to the way that the rules are ranked under the notion of fairness. Supposing, however, that the reason we should not want to make all individuals worse off has to do with the individuals' actual well-being, this view does not affect the importance of our conclusion.

about what might prove to be practically convenient for the policy maker, not a basis for endowing notions of fairness with independent normative importance.

Third, society often finds it important to instill principles of fairness—to teach *common morality*—because they are very useful rules for guiding individual behavior.<sup>26</sup> This is because individuals' ability to determine truly optimal actions frequently is limited and because individuals often will be led to act properly (rather than opportunistically) if they are rewarded by feelings of virtue and if they would be punished were they to act badly by feelings of guilt and by concerns about the imposition of social sanctions for improper behavior. The rules that prove to be most suitable for guiding individual behavior, however, should not displace proper criteria for normative assessment in the qualitatively different context of designing legal institutions.<sup>27</sup>

## 6. Conclusion

In this article, we have demonstrated that there is a fundamental conflict between notions of fairness and the Pareto principle. That is, as a matter of logical consistency, a person who embraces a notion of fairness must on some occasions favor adopting legal rules that would make every person worse off. Of course, reason alone does not tell us that it is wrong to choose legal rules that hurt everyone. Nevertheless, the conflict with the Pareto principle poses a serious challenge to proponents of notions of fairness who also care about the well-being of members of society.

## References

Benn, Stanley I. 1967. "Punishment," in Paul Edwards, ed., *The Encyclopedia of Philosophy*. Vol. 7. New York: Macmillan.

---

26. Analysts and policy makers who find notions of fairness to be intuitively appealing should display self-awareness in the light of the fact that they themselves are socialized members of society who have had notions of fairness instilled in them.

27. Indeed, as Mill (1863) and Hare (1981) emphasize, rules of common morality are themselves among the social institutions that we must be able to assess, and this requires that we first identify our ultimate normative principles. Viewing rules of common morality as those ultimate principles involves a category mistake.

- Coleman, Jules L. 1992. *Risks and Wrongs*. Cambridge: Cambridge University Press.
- Hare, R. M. 1981. *Moral Thinking: Its Level, Method, and Point*. Oxford: Oxford University Press.
- Kaplow, Louis, and Steven Shavell. 1999. "Any Non-Individualistic Social Welfare Function Violates the Pareto Principle," John M. Olin Center for Law, Economics, and Business, Harvard Law School, Discussion Paper No. 253.
- Kaplow, Louis, and Steven Shavell. Forthcoming. *Fairness versus Welfare Economics in Normative Analysis of Law*.
- Mill, J. S. 1863. *Utilitarianism*. London: Parker, Son and Bourn.
- Ross, W. D. 1930. *The Right and the Good*. Oxford: Oxford University Press.
- Sen, Amartya. 1970. "The Impossibility of a Paretian Liberal," 78 *Journal of Political Economy* 152–57.