Northwestern University School of Law
Public Law and Legal Theory Papers

Why Be Fair When You Can Have Welfare?

Anthony D'Amato*

*Northwestern University School of Law, a-damato@law.northwestern.edu

Why Be Fair When You Can Have Welfare?

Anthony D'Amato

Abstract

The Kaplow/Shavell thesis can be simply stated: If courts or legislatures pursue any value other than general public welfare, then welfare is to that extent diminished. Justice is one of those values that detract from welfare. But so too is injustice. Hence their entire thesis is a tautology.

WHY BE FAIR WHEN YOU CAN HAVE WELFARE?

Anthony D'Amato

AUTHOR'S NOTE: *I wrote the following review of Fairness v. Welfare in April 2001, based on Louis Kaplow and Steven Shavell's book-length article in the Harvard Law Review. Since then, the book itself has been published, and there have been several reviews independently reaching some of the same conclusions I've reached. However, the following text has not been changed..*

In a preprint of their forthcoming book Fairness v. Welfare, Louis Kaplow and Steven Shavell assert that any judicial decision or legislative enactment is a mistake if it is not based on the goal of making every person better off. [1] Hence every case not explainable by Paretan economic theory is in their view wrongly decided no matter how "fair" it is to the parties. A judicial decision based on fairness suboptimal because it may make at least one person worse off and could even make everyone worse off. The thesis is a shrill one, aided by the style of reader intimidation that seems to have recently emerged as the result of combining the presumptuous didacticism of law reviews with the pile-driver prose of economic analysis.

The Kaplow/Shavell thesis may be summarized as follows. Let W stand for a policy that has as its goal an increase in everyone's well being—which, of course, is the Pareto principle. Thus W is the optimal policy; a person will adopt W if she wants to increase the welfare of every individual in society. Next, let X be some policy other than W. Since X is different from W by definition,[2] then X must be less than optimal, also by definition. Therefore under X, at least one person will be made worse off. This entails the further proposition that a circumstance can exist in which adopting X will make everyone worse off.[3] In this way Kaplow/Shavell

---

[1] Louis Kaplow & Steven Shavell, *Fairness Versus Welfare*, 114 Harv. L. Rev. 961 (2001) [hereinafter *Fairness v. Welfare*]. *See also* Louis Kaplow & Steven Shavell, *The Conflict Between Notions of Fairness and the Pareto Principle*, 1 Am. L. & Econ. Rev. 63 (1999); Louis Kaplow & Steven Shavell, *Notions of Fairness Versus the Pareto Principle: On the Role of Logical Consistency*, 110 Yale L.J. 237 (2000); Louis Kaplow & Steven Shavell, *Any Non-Individualistic Social Welfare Function Violates the Pareto Principle,* John L. Olin Center paper (1999)

[2] As Kaplow & Shavell define it, "our definition of fairness includes all principles—but only those principles—that give weight to factors that are independent of individuals' well-being." *Fairness v. Welfare*, *supra* n. 364, at 989.

[3] Although Kaplow & Shavell offer a formal proof for this proposition (Louis Kaplow & Steven Shavell, *Any Non-Individualistic Social Welfare Function Violates the Pareto Principle,* John L. Olin Center paper (1999)), its truth is intuitively obvious. Briefly, 'some' is closer to 'all' than it is to 'none.' For example, if we say that there is at least one unicorn in Montana, then it is just a matter of contingent truth whether there

reach their conclusion: that any policy other than W is inferior to W in the sense that it will make at least one person, perhaps some, and perhaps everyone, worse off. If we want X to stand for Fairness, for example, then adopting a policy that is fair will lead to the unwelcome consequence that some people, or maybe everyone, will be made worse off by X. Hence, considerations of fairness should not influence the deciding of cases or the enactment of statutes.

The Kaplow/Shavell welfare thesis is of course a version of utilitarianism. But the authors think they have avoided the classic 'scapegoat objection' to utilitarianism by requiring that the policy should be aimed at making everybody better off. Under the classic 'Lee Harvey Oswald' example of the scapegoat objection to utilitarianism, a person is framed for a sensational murder in order to quiet the social unrest that might otherwise follow the public's realization that the killer is still at large. While that might make everyone *else* better off, it would make at least one person—Oswald—distinctly worse off. However, we can tinker with the scapegoat objection a bit in order to apply it to the Kaplow/Shavell theory. Suppose the authorities need someone upon whom to pin the murder. They hire a starving actor, offer him a lot of money, walk him through all the secret details about the murder that are known only to the police and the murderer himself, and put him on trial. Perhaps, if he is a good actor, he will break down at the end of a grueling cross-examination and admit his guilt. He is convicted and sentenced to death. He is strapped to the chair where he proceeds to deliver the most electrifying performance of his career. As soon as the witnesses have left in the belief that he is dead, he is unstrapped from the chair, led out the back door, and handed a cashier's check and a one-way ticket to Bangkok. Under this scenario, everybody is made better off. The police are happy that their professionalism has been vindicated—they quickly caught the killer with evidence beyond a reasonable doubt. The family of the victim feels vindicated—they believe the story. The real killer is delighted that he has escaped and now has total immunity (the law enforcement authorities could never afford to arrest him even if an eyewitness showed up). The actor is happy—he never needs to work again. The public is happy in their belief that justice has been done. Even Vilfredo Pareto is happy in that big welfare state in the sky. Hence, the Kaplow/Shavell welfare model, although ensuring that *everyone* be made better off, has not avoided the obvious unfairness associated with the scapegoat objection.

The other classic problem with utilitarianism concerns the computation of the greatest good for the greatest number. Utilitarianism theory does not solve this

---

are many unicorns in Montana or even whether all the unicorns in the world are in Montana. But if we say there is no unicorn in Montana, we are *infinitely* distant from the proposition that there is at least one unicorn in Montana. Accordingly, if W is the optimal policy that makes no one worse off, any other policy X_i is 'closer' to X_j than it is to W, where $_i$ and $_j$ can take on any value from 'one person being worse off' to 'all people being worse off.' However, it should be noted, without going into great detail here, that the authors enlist this formal logical result to argue that even in cases where W and X lead to the same result, W is preferable because X *could* lead to a result in which everyone is worse off. *Fairness v. Welfare, supra* n. 364, at 1012 n102. The authors do not reveal the mechanism (is it intrinsic or extrinsic, for example?) that would explain why X could lead to such a result..

problem; rather, it is left up to the wise determination of an impartial spectator.[4] No one knows what the qualifications for spectatorship might be, or what instructions to give the spectator.[5] Kaplow and Shavell begin their argument by an admission:

> [A] method of aggregation is of necessity an element of welfare economics, and value judgments are involved in aggregating different individuals' well-being into a single measure of social welfare.

So far so good, though perhaps they should have voided their own thesis by saying, more accurately, "*unfathomable* value judgments." Nevertheless, they proceed to make their critical move:

> Various methods of aggregation are possible. For example, under the utilitarian approach, social welfare is taken to be the sum of individuals' utilities. Alternatively, the well being of worse-off individuals might be given additional weight, as under the approach associated with John Rawls, wherein social welfare corresponds to the utility of the worst-off individuals. In this Article, we do not defend any specific way of aggregating individuals' well-being; that is, we do not endorse any particular view about the proper distribution of well-being or income.[6]

The most remarkable word—truly a howler—in this remarkable paragraph is the word "defend." Imagine a 400-page proof in a mathematics journal of Fermat's Last Theorem in which one of the steps somewhere in the middle of the proof says 'Now we square the circle. We do not *defend* any specific way of squaring it.' In the same way, Kaplow & Shavell would have us take on faith the existence of a method of aggregation that they lack the space to defend, considering that they were only allocated a total of 427 law-review pages (a number that gives a wholly new meaning to the term 'article').

Kaplow & Shavell know that the interpersonal comparison of utilities is impossible. But they want us to gloss over this inconvenient fact. They want us to assume that even though a judge cannot measure, quantify or aggregate the utilities of the two parties standing right in front of him, it is child's play for the judge to measure, quantify or aggregate the utilities of 284,000,000 people in the United States.[7] As we have just seen, Kaplow & Shavell tell us that "various methods of aggregation are possible." Such as what? The first 'method' as quoted above is: "under the utilitarian approach, social welfare is taken to be the sum of individuals' utilities." This is a *method?* It is nothing more than the very conclusion they're in

---

[4] The best brief discussion of the role of the spectator in utilitarianism that I have seen is John Rawls, A Theory of Justice 19-24 (rev. ed. 1999).

[5] All we know is that there are at least two qualified spectators: Kaplow and Shavell.

[6] *Fairness v. Welfare*, at 987-88.

[7] The resident population of the United States as of March 1, 2001, is 283,707,367. *See* http://www.census.gov/cgi-bin/popclock. The Kaplow & Shavell approach to aggregation reminds me of the story of the accountant who reports to the CEO that the company will lose money on every single widget that they decide to produce. The CEO replies: "That's OK, we'll make it up in volume."

the process of begging. Well, let us proceed. They say they have "various methods." The word 'various' suggests several more methods. However, all they mention is *one* more method. Considering the fact that the first method they cited wasn't a method at all—summing utilities is the problem, not the solution—this next one is the *sole* method that they provide to the reader. It had better be good. The method, as above quoted, is that "the well-being of worse-off individuals might be given additional weight, as under the approach associated with John Rawls, wherein social welfare corresponds to the utility of the worst-off individuals." But Rawls never said anything like that. Rawls simply wanted to transfer *money* from the best-off to the worst-off, defining the best-off as those who have the most money and the worst-off as those who have the least money (a perfectly quantifiable determination). Rawlsian transfer payments have nothing to do with the utility of individuals. For example, under Rawls's system you could conceivably take money from a rich person with the result that her life is made happier and give it to a poor person making his life miserable. That result would not count against Rawls's thesis, because Rawls isn't talking about utility at all; he's talking about social fairness (which in his view relates to the just distribution of wealth).[8] When Kaplow & Shavell title their article *Fairness Versus Welfare* and try to convince us that between the two we should choose Welfare and reject Fairness, how can they possibly enlist Rawls—who would do just the opposite and choose Fairness while rejecting Welfare—as providing the sole example for their thesis of aggregating utilities? For them to say that the Rawlsian 'method' is acceptable is like saying, in the course of the proof of Fermat's Last Theorem, that one of the various acceptable methods for squaring the circle is to use straightedge and qumpass. A qumpass is a lot like a compass except it also has the ability to square circles.

Kaplow & Shavell try to persuade us that things go wrong when a policymaker's goals include fairness, justice, morality, or respect for individual rights.[9] But at the same time that they inveigh against these goals, they apparently fail to come to grips with the fact that their own thesis proves that policymakers also go wrong when they seek *un*fairness, *in*justice, *im*morality, or *disdain* for individual rights. For the selfsame 'logic' of the Kaplow & Shavell thesis that instructs policymakers to reject fairness also instructs them to reject unfairness. The term 'X' in Kaplow & Shavell's thesis can stand for *any* goal other than the goal of making every individual better off. Thus if policymakers following their prescription were to use X as standing for unfairness, plug it into the theory and find that it tells them to reject *un*fairness as suboptimal, then they would be left with fairness—the very choice the authors are complaining about. In short, the authors have managed to prove their thesis and its exact opposite at the same time. Perhaps they should change their title to reflect this fact. I suppose they would turn

---

[8] Not that I agree with Rawls; I criticized his view in one of the earliest and most ignored articles ever published about Rawls. *See* Anthony D'Amato, *International Law and Rawls' Theory of Justice*, 5 Denver J. Int'l L. 525 (1975),

[9] They say: "Although 'fairness,' 'justice,' 'rights,' and similar terms are often used in different senses, it will not be necessary to distinguish among them for our purposes." *Fairness v. Welfare*, at 969 n.7.

down the suggestion *Unfairness Versus Welfare* as something that sounds too much like a nineteenth-century socialist tract. Even less good for sales, though more accurate, would be *Fairness And Unfairness Versus Welfare*. Perhaps the optimal title that would balance the interests of the academic and commercial markets is: *Batman Meets Welfare Woman and Loses: Cases and Materials.*[10]

The reason that taking fairness into account—or if you will, unfairness—must lead to policies that fail to make everyone better off, is precisely because Kaplow/Shavell have defined it to come out that way. Since they define W as the optimal policy, any other policy X—whether X stands for fairness, unfairness, or sublime indifference—will perforce be suboptimal. After a mountain of labor (427 pages with 1,047 footnotes), Kaplow & Shavell have brought forth a definition.[11] It turns out to be exactly the same definition they started with. At least no one can say that their article lacks focus. However, one might very well ask: what happens under their own definition if in a particular case X turns out to be equivalent to W? Can we then say that X may on occasion be just as good as W? Well, if that happens, Kaplow & Shavell say in a footnote, the fairness principle "collapses" into the welfare principle, so presumably we don't need fairness.[12] I confess that when I read this footnote I was momentarily snowblinded by the chutzpah. Apparently Kaplow & Shavell's idiosyncratic variation on a theme by Pareto has already become the standard for evaluating a concept like fairness that was a mere five thousand years in the making.

If so far I have been too general in my comments on the Kaplow & Shavell thesis, I would like to close by giving a few substantive examples of where their theory might lead anyone who takes it seriously. As I type these words, two major governmental issues are dominating the media: what to do about former President Clinton's eleventh-hour pardons, and how to repay to the American people a $1.6 trillion budget surplus. The Kaplow & Shavell thesis helps us find easy solutions to everyday social issues like these. All we have to do is cleanse our minds of fairness, justice, morality, and rights, and instead think of simple and practical ways of making everybody better off.

Pardongate is the easier of the two. Here my suggestion based on the Kaplow & Shavell thesis is that presidents and governors should simply sell pardons. After all, it's their constitutional perk, and we live in a market economy. Both the pardoner and the pardonee are made better off. And so is everyone else because of the tax dollars that are saved on incarceration expenses.

Since judging cases is similarly a constitutional perk, the preceding principle is readily extendable to the judicial branch. Why shouldn't a judge be

---

[10] Helpful hint for Kaplow & Shavell: Batman's famous belt can be used as an example of an acceptable method of aggravating utilities. Ouch!

[11] But along the way Kaplow & Shavell have presented a number of interesting arguments that deserve specific attention. I have not attacked their article because I think it is trivial; on the contrary, I think it is important.

[12] *Fairness v. Welfare*, at 1010 n.98.

allowed to offer a convicted defendant the choice of twenty years in prison or a cash payment to the judge of $100,000? Both judge and defendant are made better off, and society again saves the incarceration costs. However, because of the practical problem that many criminal defendants are indigent, the Kaplow & Shavell theory should encourage judges to be more creative and consider applying the thesis to civil cases. Suppose a plaintiff who was crippled when hit by the defendant's car is awarded $10 million by a jury. On appeal, the lawyer for the defendant's insurance company is invited into the judge's chambers. "As you know," the judge says, "corrective justice is a fuzzy concept. All I care about is striking the right social balance for the future between the freedom to drive safely and the freedom to run over pedestrians. Your insured, the driver, was clearly negligent. But negligence should not necessarily be overdeterred. I'm guesstimating that the appropriate balance can be struck in this case somewhere between one and ten million dollars." The insurance attorney replies: "I know what you're implying, Your Honor, because I've been through this drill before. If you are willing to reduce the jury's verdict to one million dollars on the ground that one million strikes the social balance society needs, our company is prepared to pay you personally an extra one million. You can think of it as our punitive damages." Since the judge could have reduced the verdict to a dollar, the plaintiff is better off getting a million dollars, the defendant and judge are certainly better off, and the public feels good that the victim has received that much money. The sole problem with my scenario is that it is conducted behind closed doors. But perhaps that's only a temporary problem; if the Kaplow & Shavell thesis catches on big time, such deals between losing parties and judges could be made on the record in open court, soon to be followed by a new law-school course *Introduction to Bribery & Corruption* (part of the law-and-economics major).[13]

There is nothing wrong with selling verdicts and pardons because Kaplow/Shavell instruct us that wrong is the wrong word—we go wrong when we think in terms of wrong. Pursuing a policy of fairness can only get you into a mess, and then you'll have to read 427 pages to get out of it. Policymakers should give up all thoughts of being fair to the people they rule. The Kaplow/Shavell thesis may strike some folks as new and provocative, but I confess to a nagging sense of déjà vu. Something about the decline and fall of the Roman Empire comes to mind. The authors inform us on their first page: "This Article will subsequently appear in book form." When their article does appear in the form of a book, I hope

---

[13] Here's a Posnerian-type suggestion for class discussion in this new course: Would it make sense for juries to begin awarding damages of one penny in tort cases? Since there is no room between one penny and zero, there will be nothing for the judge to skim. The judge will be forced to *raise* the jury's damage award in order to bring it up to the social balancing point. Since this judicial increase would be entirely unwelcome to the defendant, the defendant would have no incentive to pay any additional money in bribery charges. (Why?) Suppose the judge retaliates by threatening to add $X to the amount by which he will raise the jury's verdict and then offers the defendant a deal to lower it by $Y (which is less than $X) if the defendant pays an amount of one-half $Y to the judge directly. If the defendant refuses this offer, what can the judge do? If the judge makes good on his threat, then won't the plaintiff demand the full amount, leaving the judge nothing? I offer this suggestion out of a suboptimal sense of fairness to Kaplow & Shavell so that their thesis might not lead to quite the intellectually stultifying consequence for law-school teaching than it may otherwise appear to require.

for the sake of making Kaplow & Shavell better off that a bold announcement will appear across the front cover: Soon to Be a Major Motion Picture.[14]

The issue of the $1.6 trillion budget surplus allows us to look more deeply into what Kaplow/Shavell mean by fairness. But first for my solution: I propose that the problem of the budget surplus can be solved by drawing up a list of the 1,500 most influential political officials in the country—the President, the Cabinet officers, all Senators and Representatives, all the judges on the Supreme Court and the federal courts of appeals, all the state governors and lieutenant governors, and political hangers-on, stopping at 1,500. Then pay each of these persons one billion dollars in cash. Deducting the total of these payments from the budget surplus of $1.6 trillion leaves us with a remainder of $100 billion. Since we have to make *everyone* in the United States better off, we do so by paying the remainder out in $350 checks to each man, woman, and child. My scheme is of course pragmatic: it should not be too hard convincing 1,500 politicians that this plan comports with their rational self-interest, so long as Kaplow/Shavell have convinced the American people that any policy based on fairness is public enemy number one.

However, Kaplow/Shavell could reply to the budget-surplus example as follows: an individual recipient of a $350 check might have a *taste* for fairness that is worth more to her than $350.[15] If she finds my budget-surplus distribution scheme *that* unfair, then she has not been made better off, and so my scheme is defeated. At first blush, this reply's willingness to reintroduce fairness as a criterion makes it appear that the authors have kicked fairness out the front door only to allow it to reenter through the back door. What about their claim that a taste preference for fairness is not conceptual but rather empirical? As they put it, "the welfare economic significance of a notion of fairness depends directly on the strength of individuals' actual tastes for it and is thus an entirely empirical issue."[16] This claim is hollow unless the authors can propose or show a way of measuring individuals' actual tastes for fairness. They support their claim by handing the reader a promissory footnote: "We comment on ways to measure such tastes in [a later subsection of the article]."[17] Flipping some 375 pages ahead, we encounter the subsection referred to, and find that they have literally delivered on their promise: they do *comment* on ways to measure the taste for fairness. Stating that the measurements are best left to "statisticians and opinion researchers,"[18] their comments range from concluding that opinion research "will often have limited value" to concluding that polling information is notable for its "unreliability."[19] These comments are all the reader receives about "opinion researchers" who, with

---

[14] With Madonna in the role of Chairperson of the White House Council of Economic Advisers.

[15] As the authors put it: "[I]f individuals have internalized a social norm that is related to a notion of fairness, they may have a taste for fairness, in the sense that they may feel better off or worse off depending on whether their conception of fairness is reflected in legal rules or in the actual operation of the legal system." *Fairness v. Welfare*, at 974-755.

[16] *Id*. at 975.

[17] *Id*. at 975 n.24.

[18] *Id*. at 1350.

[19] *Id*. at 1351.

the exception of "statisticians," are the persons *best* able to measure tastes for fairness. As for statisticians, the authors drop them without the benefit of their comment. Perhaps no comment is possible; after all, statisticians can only come into the picture after someone else has done the individual measurements. The reader is left with the impression that this is just another example in the Kaplow & Shavell article of solving critical issues by straightedge and qumpass.

It could be quite difficult for anyone to prove that any recipient of $350 really values social fairness more than that amount. Most people will *say* they do if you ask them, but how can any opinion researcher be sure that they are providing a precisely accurate introspective measurement of their feelings? If you were to approach any number of the several million citizens who dodge their income taxes every year and tell them that you are an opinion researcher anxious to know why they behave the way they do, they are most likely to tell you that income taxes are brutally unfair and they are engaged in acts of effective civil protest. But how can we be sure empirically that it is their taste for fairness that outweighs the risk of getting caught?

An easier example to discharge *my* burden of proof on this argument is the Lee Harvey Oswald 'scapegoat' scenario I suggested earlier. Recall that this example resulted in everyone being made better off. Thus there cannot be any doubt that the example coheres with the utilitarian foundations of welfare theory without the embarrassment of the scapegoat problem. But now I want to make a small addition to that scenario to show that fairness is more fundamental than welfare and in fact trumps welfare. Suppose a few weeks later an investigative reporter reveals the police conspiracy to the public—through convincing videotape evidence including interviews with the actor in Bangkok showing a copy of the check he received for his services, and interviews of several police officials who were in on the plan and now 'regret' what they have done. Is there any doubt whatsoever of the public's reaction? Would any prominent academic or judicial utilitarian be willing to appear on one of the ubiquitous talking-head news shows to defend what the police did on the ground that it was important for the general good of society to lie to the American people?

The 'scapegoat' example is at bottom a case where *full disclosure* triumphs over Kaplow and Shavell's welfare-statism. The welfare-seeking policymaker or judicial lawmaker very often has to hide from the public the real facts and considerations that influence his decisions. There is something inherently surreptitious about making rules for the general public that violate fundamental fairness. The judicial welfarist must conceal from the litigants the fact that he is inventing new law at their expense (or at least at the expense of the designated loser) in order to provide for the general welfare (as the judge sees it).

Kaplow and Shavell have created a kind of miniature welfare world, like the water-filled globe paperweights sold at Christmas that you overturn and snowflakes descend upon a sleepy little town. Their miniature world is self-contained, self-consistent, and self-defined, but it is not the real world. Indeed, it is a *fake* world

that owes its novelty to illusion-making. The very internal consistency of their model isolates it. Their model does not 'plug in' to the real world, and hence Kaplow/Shavell must argue strenuously—but ultimately to no avail—that policymakers in the real world ought to conform themselves to the surreptitious processes of their make-believe welfare world. To the extent that they have made their argument consistent, they have made it irrelevant.