Book Review

The Grounds of Welfare

Jules L. Coleman†


Louis Kaplow and Steven Shavell are talented and distinguished legal academics who for the past several years have been working jointly on a massive project in normative law and economics. The project’s goal is to answer the question: What are the criteria by which legal policies (rules, standards, decisions, and other authoritative acts) ought to be assessed and proposals calling for their reform to be evaluated? In answering this question, they consider two normative frameworks—one defined by a concern for the impact of policies on human welfare, the other defined by a concern for various principles of fairness. Thus, the title of the book: Fairness Versus Welfare.† There is no surprise ending, as from the outset Kaplow and Shavell are clear that they judge welfare the unambiguous winner of the competition.

Previous iterations of the book have been in circulation for some time and available on the Internet.‡ In addition, Kaplow and Shavell have made

† Wesley Newcomb Hohfeld Professor of Jurisprudence, Yale Law School; Professor of Philosophy, Yale University. I want to thank Barry Adler, John Gardner, Mark Geistfeld, Michael E. Levine, Peter Schuck, Alan Schwartz, Scott Shapiro, Daniel Shaviro, and Seana Shiffrin for helpful comments on earlier drafts.
* Professor of Law, Harvard Law School.
** Samuel R. Rosenthal Professor of Law and Economics, Harvard Law School.
the rounds of law and economics workshops for several years, taking the opportunity such occasions provide to set out and defend the book’s central claims. Beyond that, the book has been the subject of numerous conferences and panels at professional meetings. It is unlikely, therefore, that many intended readers are not already familiar with its claims and the arguments marshaled on their behalf.

Even so, it is useful to distinguish among three groups of potential readers. The first two groups are the representatives of protagonists. On the one side are the deontologists—philosophers and legal theorists committed to the idea that some or other deontic considerations must play an independent role in assessing legal practice as well as calls for its reform. Along with everyone from Plato and Aristotle to Kant, Rawls, and Dworkin, Kaplow and Shavell are kind enough explicitly to include me in this group. This group is their target. As Kaplow and Shavell see it, no argument they could muster might convince the deontologists of the error of their ways, so hopelessly are the deontologists in the grip of a mistaken view. On the other side stand the fellow travelers along the law-and-economics highway. This group represents Kaplow and Shavell’s natural allies. Although the argument of the book might firm their resolve, and harden them in battles with the deontologists, it is not necessary to persuade them. The argument of the book will be lost on the first group and otiose for the second. This leaves the uncommitted law professor searching for an analytical and normative framework within which to organize her thinking and through which to sharpen her critical lens. The book is self-consciously aimed at capturing the hearts and minds of this segment of the legal academy.

It should come as something of a surprise, then, that among the most vehement critics of Kaplow and Shavell’s project are other advocates of an economic approach to the law. Whereas most deontologists are likely merely to dismiss Kaplow and Shavell as unsophisticated and their arguments as inadequately nuanced, the majority of law-and-economics scholars are anxious to dissociate themselves from a thesis they are convinced is dangerous to the cause. Why? The answer is that the book openly endorses precisely the imperialistic claims with which others have saddled the law-and-economics movement, often in an effort to discredit it

3. See KAPLOW & SHAVELL, supra note 1, at xxi-xxii (listing workshops at which the authors have presented portions of the book).
4. See id. at 79-81, 389-94.
as inadequately catholic or, in the extreme, uncivilized.\(^6\) Whereas the vast majority of law-and-economics scholars have been trying to make the case for including efficiency among the factors suitable to assessing legal reform proposals, the entire point of the Kaplow and Shavell argument is that the only considerations that can figure in a rational reform policy are those of human welfare—or efficiency properly construed.

One might suppose that any book that triggers so much fear and loathing—that sends its natural allies scampering for shelter and engenders apoplexy among its targets—has to be either really dreadful or of fundamental importance. *Fairness Versus Welfare* is neither. The book is divided into two parts of very unequal length. In the first part, the authors distinguish the two competing normative frameworks of fairness and welfare from one another and set forth the general framework by which they shall adjudicate between the two.\(^7\) In the second, and by far the longer, section of the book, they set out to make good on the strategy of evaluation by comparing fairness and welfare in a wide range of areas of the law—both private and public.\(^8\) The argument of the book requires for its success treating the two parts of the book as connected. That is because the objection to fairness is that the price of fairness is too high in terms of its likely impact on welfare, and so it is the burden of the second part to establish just how extensive those detrimental effects are likely to be.\(^9\) In this sense, the second part forms the evidentiary base for the thesis of the first part.

In fact, however, the second part of the book can stand on its own and constitutes a significant contribution to discussions of the impact on human welfare of various regimes of rules, standards, and policies in a wide range of areas of the law. The source of consternation for "friend" and foe alike is the first part of the book. Whereas the second part is nearly invaluable to anyone interested in policy analysis and legal reform, the first part's argument is entirely unsuccessful. Unfortunately, the overall argument of the book depends crucially on it.

*Fairness Versus Welfare* claims that welfare, and not fairness, is the standard appropriate to assessing the law and calls for its reform. This is a normative claim and, as such, requires normative argument on its behalf. Any suitable argument for the authors' claim then will consist in a set of


\(^7\) KAPLOW & SHAVELL, *supra* note 1, at 3-81. The book concludes with a few chapters that elaborate upon the framework presented in the first part of the book. For purposes of my bipartite division, these chapters can be treated as addenda to the first part.

\(^8\) *Id.* at 85-378.

\(^9\) See *id.* at 58 (noting that later chapters will document the extent to which fairness-based policies and rules diminish human welfare).
reasons or grounds for the claim that welfare, and not fairness, is the appropriate basis for assessing law and its reform. The burden of providing an account of what is to count as grounds or reasons for that claim is the task of the first part of the book: the evaluative framework. Sadly, instead of discharging that obligation, *Fairness Versus Welfare* serves up empty tautological claims and underdeveloped putative causal explanations—explanations, moreover, that were they in fact adequate, would be so strong as to undermine, rather than support, the book’s overall thesis. *Fairness Versus Welfare* makes a bold normative claim, but it offers no argument adequate to support it.

In Part I of this Review, I summarize the debate on the normative foundation of efficiency prior to the publication of the Kaplow and Shavell book. In Part II, I criticize Kaplow and Shavell’s argument that welfare is the uniquely appropriate standard for the assessment of the law and proposals for its reform. In Part III of this Review, I sketch an alternative account of the value of welfare. On that view, however, whatever it is about welfare that explains its value and aptness for assessing the law also explains why fairness is valuable and appropriate to assessing the law. In short, Kaplow and Shavell’s account of welfare fails to explain its value and its role in evaluating the law. On the other hand, any plausible account of welfare that is capable of explaining its value explains as well the value of fairness and its appropriateness to evaluating the law and proposals for its reform. The central claim of the book is not just inadequately defended, but, at the end of the day, unsupportable.

I. WHY EFFICIENCY? THE DEBATE PRIOR TO *FAIRNESS VERSUS WELFARE*

Law and economics has attained such a dominant position within the modern legal academy that we can be excused for forgetting how relatively young a field it is. Richard Posner’s *Economic Analysis of Law* is the work most responsible for thrusting an economic approach to law onto the broader academic landscape. The distinctive feature of that book was the claims it made on behalf of the explanatory prowess of economic efficiency. In the face of the familiar Critical Legal Studies objection that the law lacks coherence and objective, rational content, proponents of the economic analysis of law, led by Posner, argued that the law is a rational, coherent, and relatively determinate body of standards, the coherence and determinate content of which are explained by the principle of efficiency. As Posner and those who followed him argued, vast areas of the law—

especially the private law—could be rationally reconstructed as if they were
designed to produce an efficient allocation of resources.\textsuperscript{12} The claim was
not that the law should promote efficiency—only that it did.

This desire to shy away from efficiency as a normative ideal could not
be sustained, however. For law is the sort of institution that claims a
legitimate authority for itself.\textsuperscript{13} By its nature, law is coercive. Coercion is,
by definition, an interference with human autonomy and personal
prerogatives. To the extent that personal autonomy and human prerogatives
are presumptively good, coercion is presumptively bad. It requires a
defense. Law claims just such justification for itself. The claim may turn
out to be false—sometimes or often. Still, the claim is not incoherent or
necessarily false. This means that law must be the sort of thing of which the
claim could be true. Those who claim that the law is efficient also claim
that this fact about it contributes to its legitimacy. If that is so, it is natural
to ask what principles of justified political or legal authority efficiency
embodies or expresses. Posner understood the importance and
appropriateness of this line of inquiry and saw it as his burden to answer the
question: What justifies efficiency?

The burden of economic analysis is to identify a political or moral
value beyond efficiency itself that would be adequate to justify the state’s
employing its coercive machinery in order to achieve it. How could the best
interpretation of legal practice identify it as efficient if there were nothing
to be said from the moral point of view on behalf of efficiency? Law’s
efficiency might merely undermine, rather than support, its claim to
legitimacy. Nor would it be enough to associate some or other moral value
with efficiency. After all, not everything of value is justly pursued through
the law. The problem is not merely to identify some or other moral value
achieved by efficiency, but to find one that would justify promoting
efficiency through the coercive machinery of the law.

In setting out to meet this challenge, one might have expected Posner to
avail himself of the strategy of identifying economic analysis with classical
forms of utilitarianism. In promoting efficiency, the law promotes utility.
To the extent it is appropriate for law to promote utility, it is similarly
appropriate for it to seek to achieve an efficient allocation of resources.
Since Bentham at least, the claim that the law appropriately pursues utility
has an illustrious pedigree. The alliance between efficiency and utility
would have seemed natural in part because law and economics relies on the
Pareto criteria of efficiency. The Pareto criteria are themselves understood

\textsuperscript{12} POSNER, supra note 10, at 98-100.

\textsuperscript{13} This claim is often associated with Raz. See JOSEPH RAZ, THE AUTHORITY OF LAW 29-30
(1979).
in terms of the role they have played in solving problems within utilitarian moral and political theory.

We can distinguish between Pareto optimality and Pareto superiority. We begin by defining Pareto superiority and then define Pareto optimality in terms of it. A state of affairs $S$ is Pareto superior to another, $A$, if and only if no one prefers $A$ to $S$ and at least one person prefers $S$ to $A$. The notion of Pareto optimality is then defined with respect to Pareto superiority. A state of affairs $S$ is Pareto optimal provided there is no state of affairs $S'$ that is Pareto superior to it.

The Pareto rankings were introduced into the utilitarian literature in the early part of the last century in order to solve the so-called interpersonal comparability problem. They allow one to compare social states without making interpersonal comparisons of utility. If $S$ is Pareto superior to $A$, then because at least one person’s welfare or utility is improved and no one’s is reduced, a move from $A$ to $S$ increases overall utility. There is no need to make any interpersonal utility comparisons, as there would be if a move from $A$ to $S$ created both winners and losers. Thus, social scientists generally (Pareto himself was a sociologist, not an economist) took the Pareto rankings as a way of rendering claims about overall utility verifiable and thus meaningful. Given its role within both economic analysis and utilitarian moral theory, it is only natural to think, therefore, that the foundation for the economic approach to law is utilitarianism as mediated by Paretianism. The political or moral value captured by efficiency is utility, and economic analysis is part of the grand utilitarian tradition tracing itself back to Sidgwick and Bentham and beyond.

Natural as the alliance would appear, Posner would have none of it, and for the simple reason that he had been convinced by the classic objections to utilitarianism. Maximizing utility can often lead to injustice, sacrificing the one for the good of the many. Utilitarianism is a defective moral theory, an inappropriate standard on which to justify state coercion—or so Posner himself thought. If efficiency is, as he thought it was, an appropriate standard of state action, then one would have to look elsewhere to explain its normative attractiveness. To find a moral foundation suitable to efficiency, Posner looked to a particular conception of Kantian moral theory—one that emphasized the importance of individual autonomy as expressed in the capacity to consent.

14. It is worth pointing out that the so-called “interpersonal comparison” problem arose during the heyday of logical positivism, and that the problem may well be no more than an artifact of a mistaken semantic and metaphysical thesis, and not a real problem at all. At least, that is my view.

The argument he devised went as follows.\textsuperscript{16} States of affairs that are Pareto superior make no one worse off and at least one person better off. For that reason no one could object to them. In other words, everyone would agree or consent to them. States of affairs that are Pareto optimal have no states Pareto superior to them. Any movement from a Pareto optimal state will make someone worse off, and so not everyone will consent to it. People will consent to Pareto superior states and will never unanimously consent to departures from Pareto optimal states. Thus, the Pareto rankings reflect a commitment to consent and autonomy, not to utility. We need to distinguish the history of the Pareto rankings from the principles of morality to which they actually give expression. The history is utilitarian; the justification is Kantian. Or so Posner argued.

Welcome though it was, Posner’s argument generated a bevy of critical responses, including mine. In the first place, very little efficiency analysis in the law actually invokes the Pareto criteria. Most efficiency analysis relies instead on the Kaldor-Hicks criterion.\textsuperscript{17} One state of affairs, \( S \), is Kaldor-Hicks efficient to another, \( A \), if and only if the winners under \( S \) could compensate the losers such that, after compensation, no one would prefer \( A \) to \( S \) and at least one person would prefer \( S \) to \( A \). For this reason, some advocates of law and economics, like Guido Calabresi, came to refer to the Kaldor-Hicks criterion as the “potential Pareto” principle.\textsuperscript{18} States of affairs are Kaldor-Hicks efficient to others if and only if they could (were compensation actually paid) be Pareto superior. Of course, compensation is not paid, and so they are not in fact Pareto superior. That they are potentially Pareto superior has as much bearing on how they should be treated as the fact that I am potentially President of the United States has on how I should be treated now. The fact is that unlike the Pareto criteria, Kaldor-Hicks allows for both winners and losers. If the worries about interpersonal comparability are legitimate, Kaldor-Hicks reintroduces them; it does not solve them.

Those concerns are exacerbated by the fact that the Kaldor-Hicks criterion is subject to the Scitovsky Paradox. Scitovsky showed that two states of affairs can be Kaldor-Hicks efficient to one another.\textsuperscript{19} This means that Kaldor-Hicks is not even a weakly transitive ordering relationship.


\textsuperscript{18} GUIDO CALABRESI & PHILIP BOBBITT, \textit{TRAGIC CHOICES} 85-86 (1978).

\textsuperscript{19} See T. de Scitovsky, \textit{A Note on Welfare Propositions in Economics}, 9 \textit{REV. ECON. STUD.} 77 (1941); see also Coleman, \textit{Efficiency, Utility, and Wealth Maximization}, supra note 17, at 519 n.14 (providing a brief demonstration of the paradox).
Because Kaldor-Hicks does not observe transitivity, one cannot infer from the fact that $S$ is Kaldor-Hicks efficient to $A$ that $S$ has more utility than $A$. But if Kaldor-Hicks cannot be defended on the grounds that it embodies utility, it certainly cannot be defended on Kantian grounds. For there are losers under Kaldor-Hicks efficient states of affairs, and one cannot infer their consent to being made a loser." My first set of objections to Posner's argument, then, had three elements. First, efficiency analysis in the law invokes Kaldor-Hicks and not Pareto. Second, because Kaldor-Hicks is intransitive, it cannot reliably track utility and cannot be defended on utilitarian grounds. Third, if we assume that losers do not consent to their losses, Kaldor-Hicks cannot be defended on Posner's conception of Kantian grounds either.

It would be a mistake to think that these problems are reserved only for Kaldor-Hicks, for even were economic analysis restricted to the Pareto criteria, there is no Kantian, autonomy, or consent argument in the offing. Consider first the criterion of Pareto optimality. A Pareto optimal state is one that has no states Pareto superior to it. Any movement from a Pareto optimal state will produce losers. So it may be reasonable to assume that movement from a Pareto optimal state would not be unanimously agreed to. But it does not mean that movement to a Pareto optimal state from a prior state would be consented to, nor does it mean that everyone would consent to the Pareto optimal state in which they find themselves. A Pareto optimal state can itself be the result of a Pareto noncomparable change, one that produces winners as well as losers. If we presume that the losers in going to a Pareto optimal state would not consent to the move, then the move to a Pareto optimal state would not have been consented to.

A simple example illustrates the general point. At time $t$, everyone has nine units of $X$ each except Jones, who has one unit of $X$. At $t+1$, Jones has one hundred units of $X$ and everyone else has only one unit of $X$ each. Any

20. Coleman, Efficiency, Utility, and Wealth Maximization, supra note 17, at 533-39. On the other hand, if we ask whether individuals would choose to have policy made according to Kaldor-Hicks, then the answer might be "Yes" under well-defined circumstances, but these are the same conditions that Harsanyi showed would lead individuals to adopt average utilitarianism with interpersonal utility comparisons. See John C. Harsanyi, Cardinal Utility in Welfare Economics and the Theory of Risk-Bearing, 61 J. POL. ECON. 434 (1953). In that case, efficiency analysis is just a (perhaps) mathematically more sophisticated way of representing a commitment to average utilitarianism. But the defense of utilitarianism or efficiency (understood as Kaldor-Hicks) is not in terms of any particular moral value. Rather, it is simply a logical consequence of the notion of rationality defined in a particular way. Average utility falls out of our notion of rational choice and risk neutrality. Rational, risk-neutral parties will prefer or choose (the same thing on this theory, since to prefer is to be disposed to choose under appropriate circumstances) a principle of average utility as a way of distributing resources among themselves. This is what Harsanyi proves, and as David Gauthier once remarked, one does not argue against theorems. What one does is show that this is not a defense of the moral attractiveness of utilitarianism, but a consequence of a certain conception of rationality (in conjunction with the formulation of a particular choice problem). One would then have to show what moral value, if any, is embodied in this particular conception of rationality, and so on.
move from $t+1$ will make Jones worse off and thus the allocation at $t+1$ is Pareto optimal. By the same token, the move to the situation at $t+1$ would not have been consented to by anyone other than Jones. That does not mean that $t+1$ is not Pareto optimal. It is. It’s just that the Pareto optimality of the world at $t+1$ tells us nothing about whether or not it is or would be consented to.

If we move on to consider Pareto superiority, we locate the real problem with Posner’s defense of efficiency. Posner’s thought is that because no one is made worse off under a Pareto improvement and at least one person is made better off, it follows that no one would object to—or, more strongly, that everyone would consent to—Pareto improvements. But this is either false or a logical consequence of the definition of the notion of preference. In neither case can the fact that parties would prefer this or that be a reason or ground for their consent. The argument after all is simply this: People prefer $S$ to $A$; therefore, people would consent to $S$ (over $A$). The first clause is supposed to represent the idea of Pareto superiority, the second that of consent. The second is thus the grounds of the first. In fact, people sometimes choose to do what they do not prefer to do, and do not do what they would otherwise prefer to do, often because they think it wrong to act as they would otherwise prefer. So we cannot infer choice from preference. We could of course infer choice (or consent) from preference, but only if we build the notion of choice into the definition of what it is to have a preference. Often that is in fact what we do. To say that $S$ prefers $A$ is just to say that $S$ has a disposition to choose $A$ when the option is available. But when we do that, we cannot employ the notion of choice as an independent moral basis for the Pareto ranking. Quite the contrary in fact. We are merely defining the Pareto rankings in terms of hypothetical choices. In other words, the consent argument for Pareto superiority either fails or is best understood as a definition of rational self-interest.

To sum up: (1) Kaldor-Hicks, and not the Pareto criteria, is the basic standard of efficiency in law and economics. The Kaldor-Hicks criterion is intransitive. Two states of affairs can be Kaldor-Hicks efficient to one another. Utility observes transitivity, but Kaldor-Hicks efficiency does not. This suggests that Kaldor-Hicks does not embody or express the utilitarian ideal. (2) States of affairs that satisfy the Kaldor-Hicks standard may produce losers as well as winners. The losers cannot be expected to consent to their losses, or at least we cannot infer that they will. Therefore, there is no Kantian or consent defense for Kaldor-Hicks efficiency. (3) Nor is there a consent-based defense of Pareto optimality in the offing. On the assumption that losers will not consent to their losses, all we can say is that once at a Pareto optimal point, individuals will not unanimously consent to departures from it. (4) Nor can one infer that Pareto superior states are consented to. One can infer that Pareto superior states are preferred to those
states Pareto inferior to them. But the fact that they are preferred does not entail that they are consented to, unless preference is defined in terms of consent. In that case, the claim that Pareto superior states are consented to expresses a definition, and thus consent cannot ground or justify Pareto superiority, being completely constitutive of it. Or so I have argued.

II. FAIRNESS VERSUS WELFARE:
ASSESSING THE KAPLOW-SHAVELL ARGUMENT

This is the backdrop against which we approach our discussion of the Kaplow-Shavell book—a book that explicitly limits itself to addressing only these and other issues in normative law and economics. This is not a book that extols the explanatory virtues of efficiency or the importance of modeling legal problems as ones about the efficient allocation of resources. It is a book whose central claim is that considerations of welfare are the only defensible grounds on which to assess legal policy and proposals calling for legal reform. This is a claim that presupposes the value of efficiency, that invites us to reconsider the very same questions that Posner and his critics took up twenty years ago. At the end of the day, after all, the book’s claim is interesting only if both fairness and welfare are at least prima facie plausible candidates for assessing legal practice. Were fairness not even a plausible candidate for assessing legal practice, the claim that welfare is more appropriate to the evaluation of the law than is fairness would be both unimportant and uninteresting. Were welfare an implausible candidate for assessing legal practice, the claim that it is more appropriate than fairness in evaluating law would be no more than a bad joke.

This means that several burdens fall to Kaplow and Shavell. In the first place, because the aim of the book is to compare welfare and fairness with regard to their value as standards for evaluating the law, they owe us accounts of welfare and fairness.

21. See KAPLOW & SHAVELL, supra note 1, at 4 n.3.
22. See id. at 3 (“Our central claim is that the welfare-based normative approach should be exclusively employed in evaluating legal rules.”); see also id. at 5 (“The design of the legal system should depend solely on concerns for human welfare.”).
23. In a book of this length, a reviewer is likely to find much with which to take issue. Indeed, I found no shortage of such disagreements, but I want to limit my discussion to this, the central argument of the book. I cannot resist, however, pointing out that the authors often show themselves incapable of taking on the issues they tackle on the philosophical grounds they have chosen. One place where their lack of philosophical understanding is especially noteworthy is right at the core of the book—otherwise I would be inclined just to let the issue pass.

One of their persistent criticisms of deontologists is that we do not seem capable of settling on a shared definition of various of the notions of fairness at play. See, e.g., id. at 45-47, 86-99. We do not, for example, agree about what corrective justice is or about the nature of retributivism or distributive justice. This criticism is doubly mistaken. Most importantly, philosophers of law are not in the business of defining “fairness” or cognate terms like “corrective justice” and
We need to know what it is we are comparing. It would be demanding too much to require that Kaplow and Shavell defend one or another conception of fairness and welfare as uniquely correct or better than a range of plausible alternatives. Whereas Kaplow and Shavell need not defend accounts of welfare and fairness as correct, they do need to offer accounts of each that answer to several adequacy conditions. One crucial adequacy condition is the requirement that any account offered must have the resources sufficient to explain why welfare and fairness are valuable. More than that, really. The accounts offered must have resources adequate to explain why both are apt for the assessment of law. Not every value, after all, is one suitably pursued through the coercive machinery of the law. An account of welfare that left it mysterious why a rational policymaker might think that legal policies ought to be assessed according to their impact on welfare would fail as an account of welfare. Similarly, an account of the nature of fairness that left it mysterious why a rational policymaker might argue that the law ought to conform to the demands of fairness would fail as an account of fairness. Even if it is too strong an adequacy condition to impose on an account of either welfare or fairness that it be capable of explaining the aptness of either for assessing law, the condition is minimally necessary to make the Kaplow-Shavell book interesting. After all, their explicit aim is to show that welfare is a more appropriate criterion for assessing legal policy than is fairness, and that project is interesting only insofar as fairness and welfare are both plausible criteria for assessing legal policy in the first place.

Beyond that, Kaplow and Shavell must provide a standard for deciding between the two. They need to defend that standard as appropriate and argue that applying it to the relevant facts leads to the conclusion that welfare is uniquely apt to the assessment of legal policy. Focusing on the standard itself for a moment, it is important to note that it might take a broader or a narrower scope. Someone might defend the unique appropriateness of welfare (or fairness) as a tool for assessing legal practice by showing first that welfare (or fairness) is the correct standard for assessing all human action. Or one might argue that welfare (or fairness) is uniquely appropriate to assessing the law while setting to one side concerns

"retributivism." We are not providing a semantic or meta-semantic account of terms, but a theoretical account of the nature of the thing to which the term arguably refers. We disagree with one another about what fairness is, what corrective and distributive justice are, and, indeed, what welfare is. Ours is not a disagreement in the first instance as to the semantic content of "fairness" or "corrective justice," for example. In claiming that fairness is appropriate to the assessment of law, we cannot be understood as making the claim that the content of "fairness" is appropriate to assessing law. Our disagreements are theoretical, not semantic or meta-semantic (although as philosophers of language, we may—and do—have such disagreements as well). And once we realize that our disagreements are theoretical, not semantic, it is hardly surprising that we disagree. After all, the content of political principles—and the demands they impose—are nothing if not controversial.
about which norms are appropriate to the assessment of human or political actions more broadly.

Most, but not necessarily all, deontologists adopt the view that the standards appropriate to assessing political or legal action need not apply to human conduct more generally. This means that the aptness, say, of corrective or retributive justice for assessing tort and criminal law respectively is not in general thought to depend on whether compensatory or punitive practices within the family are similarly regulated by principles of corrective and retributive justice. In contrast, utilitarians incline to the view that the principle of utility is appropriate for the assessment of legal or political action just because all action is appropriately assessed by its impact on utility. The principle plays out differently in different contexts, but it remains the appropriate ultimate standard of assessment in all.

Because of this difference in scope of application, the standard for assessing the appropriateness of fairness and welfare as normative frameworks for the law has to be tailored to law. A welfarist or utilitarian is free to believe and contend that welfare or utility is uniquely suitable to assessing human conduct broadly, but she cannot count it against her deontologist rivals that retributive or corrective justice is not. She cannot, that is, unless she is also prepared to offer an additional argument to the effect that a norm is appropriate to assessing legal practice only if it is appropriate to assessing human conduct more generally—only, in other words, if she is prepared to argue that the political must be derivable from the ethical. Kaplow and Shavell offer no such argument, nor do I have reason to think that they would be inclined to do so. Thus, they must take on the deontologist on the narrower ground that welfare is superior to fairness as a criterion for assessing legal practice, setting to one side the relative merits of both in assessing human conduct more generally.

To sum up to this point: To support the claim the book makes, Kaplow and Shavell must first provide accounts of welfare and fairness that explain why both are apt for the assessment of legal policy. Then they need both to identify a standard for choosing between the two and to defend its appropriateness. Finally, they need to argue—on the basis of relevant facts about law, fairness, and welfare, together with the relevant evaluative standard—that welfare is uniquely appropriate to the assessment of law. This is the kind of argument the central claim of the book demands. The most striking feature of the book is that there are no such arguments in it. There are no explicit substantive accounts of either welfare or fairness

24. It is worth noting, as Rawls has, that the classical utilitarians were concerned primarily with questions of institutional design and not with human conduct more broadly or with specific details of legal practice. See JOHN RAWLS, A THEORY OF JUSTICE 22 (1971). Concern for the role the principle of utility plays in answering every minute detail of legal practice is very much a modern phenomenon, and not a particularly attractive one either.
offered, no argument presented that explains why either is valuable or apt for assessing the law. Nor is a standard for choosing between the two articulated, let alone defended as correct, and so there is no argument from such a standard to the book’s central conclusion. What then is there?

There is instead the following—what I will refer to as the "main argument":

(1) A person’s welfare is a function of what he or she values.

(2) To say that a person values something is to say that it can be represented as an argument in his utility function, or that it is the logical object of one of his preferences.

(3) Fairness is thought to be valuable.

(4) To say that fairness is valuable is ambiguous between the claim that (a) fairness is something that persons (some or all) prefer, and the claim that (b) fairness is valuable apart from whether or not persons prefer it.

(5) If fairness is valuable insofar as it is the object of a preference, then fairness is a constituent of a person’s welfare.

(6) If fairness is a constituent of welfare, then pursuing fairness improves welfare or is in any event compatible with welfare.

(7) If, however, the value of fairness is independent of its being preferred or valued by someone, then pursuing fairness is incompatible with welfare maximization.

(8) Therefore, whether pursuing fairness is compatible with welfare depends on whether it is an independent value. Fairness as an independent value is incompatible with welfare because it diminishes welfare.

25. KAPLOW & SHAVELL, supra note 1, at 18. The reader should note that what I refer to as the "main argument" synthesizes claims that the authors make at various points in the first part of the book; Kaplow and Shavell themselves do not lay out their contentions in so systematic a form.

26. See id. at 18 n.6.

27. See, e.g., id. at 10 (recognizing that "notions of fairness are . . . widely employed and respected").

28. See id. at 11-12, 21-23.

29. See id. at 21 (noting that individuals can have a taste for fairness, just as they can have a taste for fine wine).

30. See id. ("[When an individual has a taste for fairness], satisfying the principle of fairness enhances the individual's well-being . . .").

31. See id. at 52, 58.

32. See id.
(9) Therefore, understood as an independent value—that is, something whose value is independent of whether it is the object of anyone’s desire or preference—fairness is inappropriate as a standard of assessment.\(^3\)

(10) Because fairness so conceived is an inappropriate standard for assessing conduct of any sort, it is inappropriate to assessing the law.

This has the form of an argument, but it may be a mere tautology. Insofar as fairness is valuable as the object of desire, pursuing it is compatible with welfare. To the extent it is valuable apart from anyone’s preference for it, pursuing it is counter to preference satisfaction, and to the extent that preference satisfaction is constitutive of welfare, it is incompatible with welfare. This is no more than a tautology, and, remarkably, Kaplow and Shavell admit as much.\(^4\) Nor is it an informative tautology. It is not, in other words, a truth whose existence or import is revealed only upon seeing the connections brought to light by the argument. It simply follows from the view of welfare as constituted by the objects of one’s preferences and of fairness as the remainder, that is, as logically independent of one’s preferences, that if we pursue the latter we do so at the expense of the former. No elaborate argument is needed to support that claim or to have its insights revealed to us.

Because the argument does no more than reveal the analytic relationship between conceptions of fairness and welfare, whatever claims it makes about how fairness decreases welfare can be recast as claims about how welfare diminishes fairness.\(^5\) Thus, any conclusion about the relative appropriateness of welfare as against fairness for assessing the law can be recast as a conclusion about how fairness is more appropriate than welfare. In fact, neither conclusion would be warranted by the argument. The argument merely elaborates a tautology, and no normative conclusion follows from a tautology. One could conclude from this argument that welfare is preferable to fairness as a standard for assessing law only if one could also conclude that fairness is preferable to welfare—such is the nature of the tautology. Either conclusion is simply a non sequitur; a fortiori, so is the conclusion Kaplow and Shavell draw. So much for the main argument.

---

33. See id. at 56.
34. Id. at 7, 58.
35. Perhaps the point is better put by claiming that welfare precludes pursuing certain demands of fairness, rather than by claiming that welfare diminishes or decreases fairness. The latter way of casting the claim suggests that fairness, like welfare, is something that can be added up and maximized. It need not be, and probably is not.
Unfortunately, this is the only relatively explicit argument Kaplow and Shavell offer. There are, however, a variety of considerations that appear to play a significant role in their overall assessment of the case for welfare as against fairness, and if we identify and attend to them we may be able to construct another argument on behalf of their central claim. Kaplow and Shavell strongly believe that moral, political, and legal philosophers drawn to deontic considerations grossly underestimate and otherwise fail fully to appreciate the extent to which pursuing fairness can diminish welfare. So even if it is analytic that fairness decreases welfare, it is important to note just how much it does, or, more accurately, how much in principle it could. No one would endorse the pursuit of fairness were there reason to think that doing so could make everyone worse off. Yet that is precisely the sort of disaster pursuing fairness could occasion, or so they argue.

Kaplow and Shavell draw two implications from the strongly adverse effects of fairness on welfare. The first is that those who support deontic standards for assessing law incur an argumentative burden. Given that (1) welfare is appropriate to assessing legal practices and that (2) fairness can impose tremendous costs on welfare, the burden is to explain why the law ought nevertheless to conform to the demands of fairness. That is a burden that falls on the deontologist, and there is no comparable burden on the welfarist. Assuming next that (1) and (2) above are true, it is then puzzling that actual policymakers and ordinary folk, as well as political philosophers, urge that legal policy should conform to the demands of fairness. In other words, in spite of the easily demonstrated adverse effects of fairness on welfare, there remain deeply rooted deontic intuitions, whose existence calls out for explanation. It is central to Kaplow and Shavell’s thinking that they believe that they have identified just such an explanation. It is, broadly speaking, an evolutionary argument. Possessing strong deontic beliefs is evolutionarily selected for. Such beliefs contribute to human survival and thus to human welfare. Evidence, moreover, of this fact is the extent to which such intuitions are reflected in the informal norms that guide relations among us.

If we put these considerations together, we can construct what I will refer to as the “subsidiary argument”:

36. KAPLOW & SHAVELL, supra note 1, at 58 (“[W]e do not believe that the full import of fairness-based analysis for human welfare is appreciated.”).
37. Id. at 52-58.
38. See id. at 62-70.
(1) It is indeed a tautology that pursuing fairness diminishes welfare.\[39\] This is no more interesting than is the equally true claim that pursuing welfare diminishes fairness.

(2) Though (1) is a tautology, the important point is that deontologists underestimate the extent to which pursuing fairness can in fact diminish welfare. In the extreme case, pursuing fairness can make everyone worse off as judged by each person’s conception of her welfare.\[40\]

(3) This means that fairness can be very detrimental indeed to welfare.\[41\]

(4) Still, in spite of the detrimental impact of fairness on welfare, many people believe that we ought to assess policies in terms of their fairness, and not in terms of their impact on welfare.\[42\]

(5) Given what we know about the value of welfare and the detrimental impact of fairness on it, the existence and persistence of such a belief is puzzling. This puzzle begs for an explanation. The explanation can be located in the mechanisms of evolutionary biology. Having strong deontic intuitions reflected in practices has evolutionary value. It contributes to survival and to human welfare accordingly.\[43\]

(6) Thus, we can offer a reasoned explanation for the value of welfare and a causal explanation for the belief in the value of fairness. Reasoned explanations rationalize and, in doing so, justify or explain the value of welfare. In contrast, causal explanations deflate the justificatory claims made on behalf of fairness.\[44\]

(7) Therefore, because welfare is rationalizable as a standard, it is appropriate to evaluating the law. In contrast, it is the belief in the value of fairness that is explained—not by reasons, but by

\[39\] Id. at 7, 58.
\[40\] Id. at 58.
\[41\] Id. Elaborating upon this claim is, of course, the main purpose of the book's large second part.
\[42\] Id. at 62.
\[43\] See id. at 62-70.
\[44\] See id. at 62-63, 69-72, 77. For instance, Kaplow and Shavell argue as follows: [The] source of the appeal of notions of fairness—that they are associated with social norms to which we have an attachment—does not carry any implication that they should receive weight as evaluative principles when choosing legal rules. Quite the contrary is the case... If we were self-conscious about the role of social norms and the origins of our instincts and intuitions about them, we would not be led to attach independent weight to notions of fairness for the purpose of assessing legal policy. Id. at 71-72.
causes. So we must assess the law by considerations of welfare, even as we recognize the forces of nature that pull us to fairness: a pull our rational selves must resist if we are to do what is right.45

The main argument is an elaborated tautology offered in support of a non sequitur. The subsidiary argument is not a tautology, but it is no less problematic and unpersuasive. Let's begin with the thought that philosophers drawn to assessing legal practices and policies along broadly speaking deontic lines (fairness or justice) underestimate the extent to which pursuing the latter can diminish welfare. Presumably this is an empirical claim about deontologists, and, if so, it is false. Each of the traditional and widely known objections to utilitarianism put forth by deontologists presupposes that pursuing fairness or justice comes at a very high price to utility (or welfare). All the old war-horse examples—including the case of punishing the innocent—are constructed around the conflict between fairness and welfare (broadly construed). In the usual case, considerations of security and welfare make an overwhelming argument for punishing an innocent man. The question is whether we ought to. The deontologist argues we cannot. There is simply no way of understanding the deontologist's argument other than by attributing to him an implicit acknowledgment of the high cost to welfare of acting in conformity with the demands of justice. Indeed, there is no way of understanding any of the standard deontological objections to utilitarianism of this form other than as acknowledgments of the claim that conformity to justice demands much in the way of costs to welfare. So much is taken for granted. It is a further question whether the deontological arguments are always convincing—that, in other words, the costs to welfare are worth the price of justice. And it is a further question still whether deontologists must be committed to the view that no price to welfare is too high when incurring it is necessary to conform to the demands of justice. But there is simply no question that deontologists are more than adequately aware of the extent to which

45. See id. at 69-72, 77, 80-81. Kaplow and Shavell perhaps put the point most starkly in the following passage:

Id. at 80-81.
conforming to the demands of justice can come at a very high cost indeed to welfare.

Kaplow and Shavell similarly fail to appreciate that if what they offer up is a good argument, it is as telling against the welfarist as it is against the deontologist. For if conforming to justice can greatly diminish welfare, it is equally true that pursuing welfare can greatly diminish justice. The very same counterexamples to utilitarianism that presuppose the extent to which one must forgo welfare to conform to the demands of fairness can be read as indicating the extent to which pursuing welfare imposes costs on fairness. Isn’t that just the point of the punishing-the-innocent kind of example? That is, if all we have in mind is pursuing welfare, then we run the risk of imposing the greatest sorts of injustices—including punishing those we know to be innocent of wrongdoing. Similar remarks are in order for all such counterexamples to utilitarianism. In just the same way that they require the deontologist to confront the cost to welfare of conformity with justice, they force the welfarist to confront the cost to fairness of a single-minded pursuit of welfare. That is why so much ink has been spilt on these examples. They vividly raise the conflicts between two different conceptions of right action, and the costs of a single-minded devotion to either. Not only is it simply false that deontologists fail to appreciate the costs to welfare of commitment to conformity with fairness, but also, if true, the charge applies equally to the economist. We have uncovered no truth about deontology or about its relationship to welfare. All Kaplow and Shavell have provided are vivid examples of how the conflict between welfare and fairness—a conflict that is inevitable, given the way they conceptualize the two—might play out in actual practice.

The most puzzling feature of the Kaplow and Shavell argument is the apparent sense that there is something perplexing about strong deontic intuitions that calls for a causal explanation. Every component of this part of their argument is problematic—from the claim that deontic intuitions, but not welfarist ones, need explanation, to the explanation itself, to the implication that to offer a causal explanation of such intuitions is somehow to deflate their justificatory force. Beyond that, again there is the problem that they fail to see that if the arguments they marshal forward are adequate against the deontologist, they are equally compelling against the welfarist. Let’s take up these problems in turn, beginning with the suggestion that

46. For example, Kaplow and Shavell note:
Our two conclusions, about how the pursuit of notions of fairness makes individuals worse off and about the lack of affirmative warrant for using notions of fairness as evaluative principles, raise the question of why legal policy analysts (including ourselves), policymakers, and philosophers, among others, find these notions so appealing.

Id. at 62.
there is reason to think that strong deontic intuitions somehow call for explanation.

Given the value of welfare and the detrimental impact of pursuing fairness on it, Kaplow and Shavell wonder why it is that otherwise intelligent people would insist that the law should conform to the demands of fairness. The view that the law should cannot be defended by reason. The best one can hope to do is to explain why philosophers and others nevertheless insist on it. Where Kaplow and Shavell fail to see the possibility of justification, they offer an explanation instead. The belief in the value of fairness cannot be justified, but it can be explained. Indeed, the fact that it can be explained is part of what counts against the claim itself. This is a difficult idea, and some care must be taken to understand it and its significance within their overall argument.

To this end, it is helpful to distinguish between two (of many possible) views about the relationship between explanatory and normative projects. The first of these is represented by a particular form of philosophical naturalism, what we may call “replacement naturalism.” Replacement naturalism is the view that normative projects—whether in epistemology, jurisprudence, or the philosophy of mind—are hopeless. Because normative projects cannot succeed, the only projects left worth pursuing are explanatory ones.

A familiar example from epistemology illustrates the general strategy of argument. We can think of the central projects in epistemology in the following way. Begin by thinking of sensory experience as epistemic inputs and beliefs, theories, or worldviews as epistemic outputs. The project of epistemology is to determine which outputs are warranted by a given set of inputs. This is the project of identifying the criteria of warranted or justified belief. Beliefs are warranted if they are appropriately supported by the inputs. What constitutes appropriate support or justification? In this sense, traditional epistemology is “justification-centered.” Famously, Quine argued that there are no a priori discoverable rules that uniquely pick out some beliefs as warranted and others as not, given a set of epistemic inputs. Philosophy has nothing to contribute to helping us identify the norms of sound or good reasoning—the norms that, if followed, would uniquely warrant some beliefs and not others. Instead of trying to determine the norms governing good reasoning—those which, if followed, would warrant beliefs—we should study how people in fact reason. Traditional

47. See Brian Leiter, Naturalism and Naturalized Jurisprudence, in ANALYZING LAW: NEW ESSAYS IN LEGAL THEORY 79 (Brian Bix ed., 1998) (distinguishing among types of naturalism); see also JULES L. COLEMAN, THE PRACTICE OF PRINCIPLE 210-17 (2001) (assessing the prospects of a naturalized jurisprudence).

epistemology would be reduced to a chapter buried somewhere in a psychology textbook as its justificatory projects gave way to explanatory projects of cognitive psychology; thus the phrase "replacement naturalism."

From the view that explanatory inquiries substitute for normative ones, we can distinguish a family of views around the claim that causal explanations have normative consequences. There are of course many senses in which explanations are normative, and many ways in which explanations can play a role in normative arguments, none of which need be controversial. Not all the views one might have about the relationship of explanation to justification are so innocent, and in a moment I want to focus on one particularly controversial claim. According to the claim on which I want to focus, a causal explanation of why someone asserts the claim she does bears on our assessment of the merits of the claim itself. In particular, in the form of the claim upon which Kaplow and Shavell rely, causal explanations of the beliefs we have and the claims we make can deflate those beliefs and claims.

Before focusing on this claim and its centrality to their argument, it is important to contrast it a bit further with some more familiar law-and-economics attitudes regarding the place of explanatory and normative projects. Some advocates of law and economics are skeptical of the value of normative projects and are inclined instead to restrict themselves to explanatory endeavors. This makes them naturalists of a sort, self-aware or not. Others are not skeptical of normative judgments as such, but feel that the special role of economic analysis is not to advance normative judgments but to contribute to the rationality of the judgments we reach by uncovering important causal connections, for example between liability rules and accident rates. Most law and economics contributes to the wisdom of our judgments and the defensibility of our social policies in precisely this way.

Kaplow and Shavell eschew these more modest understandings of the way in which causal explanations can figure in normative arguments in favor of the much stronger claim that a causal explanation of the beliefs individuals have deflates the merits of the beliefs themselves.\textsuperscript{49} As they see it, the fact that there is a persuasive causal explanation of why many people—including philosophers, policymakers, and ordinary folk—believe that the law ought to promote fairness should figure in our assessment of the claim itself. Why else would they introduce, in the course of defending welfare against fairness, the fact that such an explanation exists?

There are two connected claims here: first, that there is a persuasive causal explanation of why individuals insist on the view that the law ought

\textsuperscript{49} I do not mean to suggest that Kaplow and Shavell, jointly or severally, do not pursue these more modest projects elsewhere, including elsewhere in this book. Rather, the point is that the central argument they make against fairness depends on this stronger thesis.
to be assessed by its conformity to the demands of fairness, and, second, that the fact that there is such an explanation counts in our assessment of the underlying merits of the claims of fairness. This latter claim is the interesting one for our purposes. There is no denying that it is bold and distinctive and, as we shall see in a moment, essential to their overall argument. Bold, distinctive, and essential to the argument it may be, but plausible and adequately defended it is not. Worse, were the claim sound, it would likely undermine rather than support their overall thesis.

Without qualification, the view that the existence of a causal explanation of the fact that someone holds or asserts a particular claim undermines the truth of the claim asserted simply cannot be sustained. Often, the best causal explanation of why someone might assert that $P$ is the truth of $P$. So it cannot be that causal explanations as such undermine the truth of the claim the assertion of which is being explained. The thesis must be narrowed or weakened. Even if most causal explanations have no bearing on the truth of the claims asserted, some explanations might. Which causal explanations of the beliefs we have go to our assessment of the merits of our beliefs, and why? Kaplow and Shavell obviously believe that evolutionary explanations of our beliefs have a bearing on the merits of the beliefs themselves. For they argue both that there is an evolutionary explanation for the fact that individuals believe in the value of fairness as appropriate to assessing the law, and that this fact cuts against the claim of fairness and indirectly supports the claim of welfare.

As they see it, evolution shows that it is to our collective advantage to believe that our affairs should be regulated by fairness, and so it should come as no surprise that individuals insist that our affairs be regulated by fairness. On their view, we insist on fairness because it is to our advantage to do so, not because our affairs really ought to be regulated by fairness. Once we are suitably attuned to the evolutionary advantages of our strong deontic intuitions—advantages evidenced, they claim, by the extent to which our informal norms and practices embody fairness\(^5\)—the grip of those principles on us is correspondingly weakened. Learning the source weakens the grip, and in this way evolutionary explanations figure in our justifiably discounting the content of the claims of fairness.

It is easy to confuse this view with a familiar and more plausible one that is common in everyday discourse. In a heated debate on tort reform, one participant advances the view that a cap should be imposed on damage awards for medical malpractice, and another replies by noting that the advocate of capped damage awards is a surgeon who has been sued successfully for medical malpractice. In so replying, she means to deflate the surgeon’s claim. Another person advances the view that rents in New

\(^5\) See Kaplow & Shavell, supra note 1, at 63-69.
York City should remain stabilized, and others dismiss her view on the grounds that she lives in a rent-stabilized apartment. In these cases, reflecting on the source of the claims, we are not surprised by the opinions both advocate, and we are properly skeptical of their motives. Even so, while we may be skeptical of the motives in each case and wary of the arguments adduced, the fact remains that both the surgeon and the apartment dweller may be right. They are interested advocates, and this may increase their burden to make a compelling case, but such a case can be made. And for the obvious reason that the causal story goes to the reliability (one way or another) of the witness, not to the truth of what is asserted.

Kaplow and Shavell are making a stronger claim. For it is not their view that once we learn the evolutionary origins of the belief in fairness that we should adopt a skeptical posture to those who advocate fairness—a posture that would incline us to demand more in the way of normative argument before we are persuaded by the truth of the underlying claim. Their view is that the evolutionary argument for fairness, by itself, is part of the argument against it and for welfare.

I have attributed this claim to Kaplow and Shavell, but they nowhere explicitly make the claim or defend it. My claim is not that Kaplow and Shavell explicitly advance this view. Rather, my claim is that we must attribute such a view to them to make their argument work. To appreciate the centrality of this claim to their overall argument, we need to retrace the argument up to the point at which the evolutionary argument for the belief in fairness is introduced.

The basic argument rests on just a few ideas. As I have reconstructed it, the basic thought is that welfare is valuable and that pursuing fairness is detrimental to it. This cannot suffice to defeat fairness for the simple reason that we can run the argument in reverse. Fairness is valuable, and pursuing welfare is detrimental to it. The charge against fairness is then strengthened with the observation that pursuing fairness can greatly diminish welfare. This is no help either, since the same is true of fairness: Pursuing welfare can greatly limit our ability to act in accordance with the demands of fairness. This is basically the argument of the book, but for two ancillary discussions—one on the evolutionary origins of the belief in fairness and another on, loosely speaking, the plasticity of our deontic intuitions.51

This means that in the absence of the evolutionary argument, Kaplow and Shavell have adduced no considerations against fairness and for welfare that could not be recast to cut the other way. Without the evolutionary argument, nothing they say cuts one way or the other. The evolutionary

---

51. These latter considerations are introduced partly to bolster the evolutionary argument and partly as an independent argument against fairness.
argument thus carries a very heavy burden. It is all that stands between their having made the case for welfare and their having made no case at all. The evolutionary argument is supposed to introduce considerations that cut against fairness and for welfare. And how can that be unless the existence of an evolutionary explanation for fairness counts against the claims of fairness and for welfare? There is no way to read the evolutionary argument in the context of the book other than as an effort to deflate the claims of fairness and thus indirectly to make the case for welfare.

With so much riding on the evolutionary argument, we should pause for a moment to identify what Kaplow and Shavell would have to demonstrate in order to make good on it. Evolutionary arguments are a species of causal explanations. Causal explanations of why someone believes or asserts what he does typically do not bear on the merits of those beliefs or assertions. So one thing Kaplow and Shavell need to do is to explain why evolutionary explanations do. More than that, of course; they would need to explain why evolutionary arguments deflate rather than bolster the merits of the beliefs whose existence they are said to explain. At the most basic level, of course, they would need to show that evolution (or something like it) actually selects for beliefs. In addition, they would have to show that evolution has selected for the belief in fairness. Even this would not be enough, for they would have to show that even though evolution selects for the belief in fairness, it does not select for the belief in welfare. For if evolution selects for the belief in welfare, the entire project would collapse.

Kaplow and Shavell meet none of these adequacy conditions. They offer no basis on which to distinguish evolutionary arguments from other causal or functional explanations—that is, no grounds for thinking that evolutionary explanations go to the underlying merits of the beliefs we have in a way in which other causal explanations do not. They offer no grounds for thinking that evolutionary arguments deflate rather than bolster the merits of our beliefs that have evolutionary origins. They provide no reason for thinking that evolutionary factors select for beliefs. If evolution selects for beliefs, they provide no reason for thinking it would select for the belief in fairness and not for the belief in welfare.

We can set aside many of these worries for now because, even on its own terms, the argument fails. It is something of an exaggeration to say that Kaplow and Shavell offer an evolutionary explanation for the belief in fairness, or for strong deontic intuitions more generally. They do little more than gesture at an evolutionary-style argument.\(^5\) Roughly, the argument is

---

52. In doing so, they do not indicate that they understand the limits within which evolutionary arguments are appropriate. Evolution does not select for particular beliefs held by particular persons. If anything, it selects for human capacities, for example the capacity to have beliefs, or to reason, etc. Too often, evolutionary arguments are really no more than metaphors for a “style” of argument. This is not the place to explore the extent to which social scientists have misunderstood
that there are evolutionary advantages to believing that human affairs ought to be regulated by considerations of fairness. Once we understand the evolutionary benefits of maintaining strong deontic intuitions, the puzzle is solved. Those who insist on evaluating the law by its conformity with the demands of fairness are advancing a view that cannot be sustained by appeal to reason or to value, but which can be explained evolutionarily—as a false but valuable one.

If Kaplow and Shavell are correct, the belief in the value of fairness (as valuable independent of its connection to welfare) is mistaken. Not just a harmless false belief either, but one that, to the extent that it has affected our legal practices, has been the source of waste, misery, and misfortune. The correct belief is in the value of welfare. The claim that evolution selects for fairness implies that evolution would select for a false belief. Other things being equal, we should want an accurate road map of the world as we try to negotiate our way through it. After all, an inaccurate road map will likely lead us astray. If we think of our beliefs as constituting a map of the world, why shouldn’t we expect evolution to select for true beliefs, rather than for false ones? Some of the beliefs on which we rely in making our way through life are normative ones—beliefs about how we ought to behave, what we owe one another, and how governments should regulate affairs among us. Shouldn’t we expect evolution to select for true normative beliefs, not for false ones?

We don’t have to claim that evolution must select for true beliefs in order to undermine Kaplow and Shavell’s argument. It is enough to note that they provide no reason for thinking that evolution would select for this particular false belief in the value of fairness as opposed to the true belief in the value of welfare. If anything, there is at least as much reason for thinking that evolution selects for welfare as for fairness. If evolution selects for beliefs, it is likely to select for beliefs that contribute to our capacity to negotiate the world successfully. In general, true rather than false beliefs are more likely to contribute to our capacity to negotiate the world. There is probably a better evolutionary argument for the economist’s insistence on welfare than there is for the philosopher’s insistence on fairness. Given their view that an evolutionary explanation of the beliefs deflates the content of the beliefs, this would not be a good outcome for them. On the one hand, evolutionary pressures explain the belief in welfare. On the other hand, they are committed to the view that evolutionary arguments deflate. The deflation claim seemed helpful when there was
reason to think that evolutionary considerations would cut against fairness. The argument looks less helpful when it is likely to cut against welfare.

These evolutionary considerations are even more troubling than we have suggested so far. If true beliefs are the ones that are in fact evolutionarily selected for and if the evolutionary argument cuts against them, then the only claims that are not undercut by evolution are the ones evolution does not select for: namely, false, unhelpful ones. Could anything be more implausible than this? In fact, all this is best taken as a reductio of the entire form of argument.

This line of objection seems so devastating to the evolutionary argument that I wonder whether I have misunderstood the role that evolutionary considerations are to play in their thinking. Perhaps the evolutionary argument is offered to make a different point, namely that everything, even the belief in fairness, survives only because it contributes to welfare. That argument might run as follows. We have three categories of beliefs on which we might act: the belief in the value of welfare; the belief in the value of fairness, where fairness is reducible to an argument in this or that person’s utility function; and the belief in fairness, where fairness is thought to be something valuable whether or not anyone has a taste for it. Acting on the belief in the value of welfare (other things being equal) promotes welfare. Acting on the belief in the value of fairness as something valuable insofar as individuals have a taste for it also promotes welfare (other things being equal). There is an apparent tension when philosophers and others argue that we should act on a belief in the value of fairness as something valuable whether or not anyone has a taste for it. Such action would appear to be inconsistent with welfare. But the evolutionary argument shows us that this inconsistency is more apparent than real. For the belief in fairness as an independent value itself survives only because it is welfare enhancing. Evolution shows us that everything converges on welfare, even fairness; everything that survives contributes to welfare.

If this is the argument Kaplow and Shavell have in mind, then rather than supporting the central thesis of the book, it overwhelms it. If everything survives because it contributes to welfare, then that’s all there is to it. It does not matter how we assess the law. If we think it appropriate to assess the law by its impact on welfare, that’s good because doing so contributes to welfare. If, on the other hand, we come to the view that we ought to assess the law by the extent to which it conforms to the demands of fairness, no problem. For we have come to that belief evolutionarily.

53. The difference between the two conceptions of fairness is this: In one sense, fairness is valuable because someone values it (as judged by the fact that he or she has a taste for it). In the other sense, fairness is valuable, and therefore individuals ought to value it—that is, they ought to prefer it or have a taste for it. It is a reason on which they ought to act.
And if we have come to that view evolutionarily, acting on such a view contributes to welfare. The problem with interpreting the evolutionary argument this way is that it makes a joke of the book's project; there is no meaningful sense in which there is a competition between fairness and welfare. The only competition is between direct and less direct methods of attaining welfare. Indeed, promoting welfare is hardly something that we attain, for at the end of the day, evolution makes it impossible for us to do anything else. Hardly the stuff of a classic work.

Kaplow and Shavell's reliance on evolutionary theory is a methodological nightmare. In the first place, the appeal to evolutionary considerations cannot be taken literally since evolution does not select for particular beliefs. They claim that because there are advantages to a belief in fairness, there must be an evolutionary-type argument for it. But they also believe that the belief in the value of fairness is false, and moreover that such a belief has many disadvantages as well. This is hardly adequate material to support an evolutionary hypothesis that would explain the persistence of a false belief.54 There is better reason to suppose that there is an evolutionary argument for the belief in the value of welfare—a belief Kaplow and Shavell insist is both true and beneficial. They provide no reason for thinking that evolutionary arguments go to the merits of the beliefs we have when other causal-functional explanations do not, and no reason for thinking that such explanations deflate rather than inflate the underlying claims. Worse, if evolutionary explanations do go to the underlying merits, that would put Kaplow and Shavell in a bind. There is an evolutionary argument for welfare, and, on their view, such an argument should undermine the claims of welfare. If, however, the evolutionary argument is introduced to show that everything reduces to welfare at the end of the day, then it proves too much.

Forgetting the methodological confusions that ravage the argument itself, and focusing only on what inferences we can draw from it, several options are available. If there is an evolutionary argument for fairness, then there is one for welfare. If evolutionary arguments deflate the normative beliefs they explain, the claim to the value of welfare as well as claims to the value of fairness are equally deflated. If evolutionary arguments are irrelevant to the content of the claims the belief in which they explain, then the evolutionary argument is a non sequitur. If the evolutionary argument

54. My claim is not that evolution selects for true beliefs. My view is that evolution does not select for beliefs in general. If it did select for beliefs, I would not claim that it selects only for true beliefs. There are of course persistent false beliefs, and it is an interesting question as to why they survive as long as they do given powerful countervailing evidence. That is an issue for those more versed in evolutionary epistemology than I. My point is the very modest one that other things being equal there is a better armchair case—and that, after all, is all that Kaplow and Shavell offer—to be made on behalf of the evolutionary advantages of the belief in welfare than there is for the belief in fairness.
shows that everything is reducible to welfare, then it makes a mockery of the book. At the end of the day, the evolutionary argument is either a non sequitur, proves nothing, or proves too much. In no case can it shoulder the burdens Kaplow and Shavell have placed on it.

It might be helpful if we synthesize all the arguments of the first part of the book.

(1) From the main argument we can derive the conclusion that pursuing fairness reduces welfare.

(2) From the subsidiary argument, we get illustrations of the extent to which it might.

(3) From the subsidiary argument, we also get a causal explanation of the belief that pursuing fairness is appropriate.

(4) From the subsidiary argument, therefore, we derive a deflationism about the normative claims for fairness.

(5) When all of these considerations are put together, we have an argument for the conclusion that pursuing welfare (as against the alternative of pursuing fairness) is uniquely appropriate to the legal domain.

The problems with the argument can be put as follows:

(1) Given how they define the terms, the claim that fairness diminishes welfare is a tautology and can be recast as the claim that welfare diminishes fairness.

(2) The subsidiary argument illustrates not just that fairness can greatly diminish welfare, but that welfare can in fact greatly diminish fairness as well.

(3) To the extent that the subsidiary argument offers an adequate causal explanation of strong deontic intuitions, the same argument applies to strong welfarist intuitions.

(4) To the extent to which evolutionary or causal explanations more generally deflate normative claims on behalf of deontic beliefs, they deflate analogous claims made on behalf of welfare.

(5) On the other hand, there is no reason to suppose that causal explanations deflate justificatory claims in either case.
The problem with *Fairness Versus Welfare* is that it makes a normative claim but offers no normative arguments to support it. Instead, Kaplow and Shavell present a mixture of tautological claims about the relationship between fairness and welfare, and putative causal explanations of the fact that individuals have strong deontic intuitions. The former are empty and cannot support a normative claim. In truth, the latter fail to rise above the level of mere speculations. Even as adequate causal explanations, such factors are irrelevant to the truth of the underlying claims on behalf of fairness. And, as I have demonstrated over and over, if the evolutionary argument is a problem for the deontologist, it is a problem for the welfarist as well. At the end of the day, of course, the existence of a possible causal-functional explanation of why we believe what we do has no bearing on the truth of what we believe.

### III. Welfare and Fairness Redux: Outlining an Alternative View

The central claim of the book is that the law should be assessed by its impact on welfare and not by its conformity to the demands of fairness. This claim is interesting only if both welfare and fairness are plausible frameworks within which to evaluate the law. An argument appropriate to sustaining the book's claim, therefore, would begin with substantive accounts of both welfare and fairness. Such accounts would be adequate insofar as each possessed resources sufficient to explain the value of welfare (and fairness) as well as the aptness of each for evaluating the law. Once adequate accounts of welfare and fairness were in place, an appropriate argument would then set forth and defend a criterion for choosing between welfare and fairness so conceived.

Kaplow and Shavell provide no criterion for choosing between fairness and welfare—which is puzzling in a book entitled *Fairness Versus Welfare*. They provide evidence that requiring the law to conform to the demands of fairness limits the extent to which it can promote welfare. Many of these very same examples, however, also illustrate the extent to which promoting welfare restricts the extent to which the law might satisfy the demands of fairness. And so the evidence they offer, while interesting in its own right, has, in the absence of a criterion for choosing between welfare and fairness, no bearing on the book's fundamental claim.

The reader should recall that it is Kaplow and Shavell, and not their deontologist target, who make the radical claim. They claim that the law should be assessed only by its impact on welfare, and never by its conformity with the demands of fairness. The deontologist need make no comparable claim. The deontologist need not, and likely does not, claim that the law ought not to be assessed by its impact on welfare. He claims only that in addition to being assessed by its impact on welfare, the law
ought to be assessed as well by the extent to which it conforms to the
demands of justice. Kaplow and Shavell reject even that modest claim. One
would think, with the stakes so high and the claims so strong, that Kaplow
and Shavell would offer and defend a criterion for making the choice—but
they do not.

Nor, should it be said, do they offer an account of fairness or of its
value. Indeed, they explicitly reject the need to provide an account of what
fairness is. Therefore, they offer no criterion for distinguishing corrective
from retributive justice, and both from distributive justice. All are lumped
together under the general rubric, fairness, whose value or aptness as a
standard for assessing the law remains completely obscure. On their view, it
is not necessary to explain what fairness is or why pursuing fairness is
valuable because their argument is that whatever fairness is and whatever is
valuable about it, pursuing it is detrimental to welfare.

If truth be told, the only serious claim they make is that fairness is
detrimental to welfare, and we have gone to great lengths to establish the
limited interest of this claim in making out the book’s fundamental
conclusion. Even so, one would think that with so much riding on the claim,
Kaplow and Shavell would go to similar lengths to spell out the nature of
welfare and to explain its value and appropriateness to assessing the law. If
it is decisive against fairness that it is detrimental to welfare, then welfare
must be pretty damn important. It may be, but whether it is depends in part
on what we take welfare to be. So we need an account of what welfare is
and an explanation of its extraordinary value.

In fact, Kaplow and Shavell offer precious little by the way of an
account of the nature of welfare. They have even less to say about the value
of welfare—apparently content to observe that no one denies that welfare is
valuable. Although Kaplow and Shavell do not offer what anyone would
regard as an account of welfare, they do offer a general picture of its
constitutive elements. There is enough in that picture for us to
differentiate it from other possible conceptions of human welfare, and to
ask whether welfare, so conceived, is something of such value that it could
provide the only standard suitable for assessing the law.

In what follows, I argue for two points. First, on any interpretation of
welfare plausibly attributed to Kaplow and Shavell, it is unclear what the
value of welfare is. It is even less clear why one would insist that welfare,
so conceived, is uniquely appropriate to assessing the law. If I am right, the
central claim of the book is entirely unsupported by their arguments.
Beyond that, I argue that any conception of welfare adequate to explain its

55. KAPLOW & SHAVELL, supra note 1, at 5 n.7, 38-39 & n.48.
56. This is no place for casual empiricism. Similar casual empiricism would ground the value
of fairness, for there are few among us, I suspect, who would deny that fairness is valuable.
57. See KAPLOW & SHAVELL, supra note 1, at 18-24, 409-36.
value would also explain the moral significance of deontic considerations and their aptness for assessing the law. Whatever it is about human welfare that makes it appropriate to assessing the law explains why assessing the law in terms of its justice and fairness is similarly appropriate. Any argument offered in defense of their central claim merely defeats it. Thus, I suggest that not only is the central claim of the book unsupported, it is unsupportable.

Let's begin with their partial conception of welfare. In their framework, the constitutive elements of a person's welfare are her preferences. A person can have welfare only insofar as she has preferences that obey a set of rationality constraints. If a person has a set of preferences over all possible social states, and these preferences obey familiar rationality constraints, then a person can have welfare; and her welfare is determined by the satisfaction of her preferences. The more her preferences are satisfied, the greater her welfare. Maximizing preference satisfaction maximizes welfare (for an agent).

The basic elements of welfare are preferences and their satisfaction. This leads to the natural identification of welfare with preference satisfaction. This identification is too quick, however, for the notion of preference satisfaction is ambiguous between a logical and a psychological sense. In the logical sense, to satisfy a preference is to realize it. To satisfy Jones's preference that $P$ is to bring $P$ about. Whether a person's preferences are satisfied in the logical sense is one thing; whether he is satisfied in the sense of experiencing pleasure, joy, happiness, or gratification as a result is another thing altogether.

We might flesh out Kaplow and Shavell's account of welfare in two distinct ways corresponding to these two notions of satisfaction. In the logical sense, a person's welfare is maximally satisfied when his desires are maximally realized. If this is what welfare is, why is welfare in this sense something of value? The value of welfare in this sense cannot be the value we associate with gratification, joy, or pleasure. Rather, the value of welfare is the value of seeing to it that people get what they want. And their getting what they want—having their desires or preferences realized—is valuable independent of what individuals want and whether getting what they want is met with pleasure, joy, or gratification, on the one hand, or consternation and regret on the other. Often we regret what we have chosen to do and what we have done, what we desire to do and the actions that

58. Id.
59. A person can have his preferences satisfied—that is, realized—after he is dead. His preference is therefore satisfied in the logical sense, but he secures no satisfaction in the psychological sense. The dead may be raised but that does not mean they can get a rise out of having their preferences satisfied in the logical sense. That, after all, is one of the unhappy consequences of being dead.
flow from those desires. What is so valuable about seeing to it that individuals get what they want—if what they prefer brings them no happiness or joy, or if what they want is bad for them and for others—that the law should promote welfare in this sense?

In contrast, if welfare is understood in terms of psychological satisfaction, then it is not obvious what the value is of satisfying preferences in the logical sense. After all, the psychological state that we desire to bring about by acting on the basis of our preferences may sometimes be achieved only by frustrating rather than by realizing our desires. By the same token, because individuals may secure gratification or satisfaction from all sorts of activities that are bad for them and for others as well, why would anyone think that maximizing satisfaction in the psychological sense should be the goal of law?

Although Kaplow and Shavell offer no general account of welfare, they do understand welfare in terms of preference satisfaction. But preference satisfaction is an ambiguous notion. There is an important difference between satisfaction in the logical and the psychological sense, and the relationship between the two notions is anything but unproblematic. Satisfying preferences in the logical sense is no guarantor of satisfying them in the psychological sense. On the other hand, satisfaction in the psychological sense can require frustrating rather than satisfying preferences in the logical sense. On neither account is the value of welfare obvious. And its unique or distinctive appropriateness for assessing the law remains mysterious.

I don’t mean to suggest that gratification or psychological satisfaction is undesirable or valueless. Nor am I suggesting that realizing one’s desires is similarly without appeal or value. Still, whatever the value of either may be, it cannot carry the normative burden with which Kaplow and Shavell saddle it. For their view is that welfare conceived in either sense is uniquely appropriate to legal policy in a way in which fairness is not.60

60. I have heard proponents of law and economics argue more than once and always with a straight face for a version of preference utilitarianism or preference welfarism based on a general normative skepticism. On such views, there are no objective values; nothing is good or bad objectively. All we have are people’s preferences or desires. The problem here is the obvious one that one cannot have it both ways. One cannot defend the normative value of satisfying preferences while at the same time rejecting the objectivity of value. What is the value, one might ask, of satisfying preferences when nothing has objective value? One cannot defend normative claims on a foundation of normative skepticism. Sometimes this claim is weakened and recast as the view that individual preferences are normatively less controversial than are the claims of justice, of what we owe one another. But this is anything but obvious. I would argue that quite the opposite is true. What is not mysterious to us is that we have no right to disregard the interests and rights of others. What is mysterious is why, in spite of this, we give such a special standing to our own interests. There may be a psychological explanation for why we accord our own interests pride of place, but what calls out for justification is the normative priority we accord our own desires. That position is, if anything is, mysterious and, in any case, quite controversial morally speaking.
This brings us to my second, and ultimately more important, point. Any plausible account of welfare that explains its value and aptness for assessing the law also explains the value of fairness and its aptness for assessing the law. We do better in understanding the nature of welfare and its value if we think of an individual's welfare not in terms of his preferences, but in terms of what is in his interests—not in terms of what he desires, but in terms of what is good for him. Among the distinguishing features of persons is that in addition to having preferences, they are planning agents. They can formulate projects and plans, invest in and execute them. They can form views about what it is they want from life, and guide their behavior by the plans they make, the agreements they reach, and the norms that regulate their affairs with others. This is part of what it is to live a life of one's own—what it is for a life to be something one does rather than something that happens to one.

Persons have interests not only in what they desire and in realizing those desires, but also in autonomy and security. They are interested not only in having their desires realized and in securing gratification and pleasure. They also have an interest in organizing a life in a way that makes sense of the desires they have. They have an interest in contributing to the way their life unfolds. The capacity to live a life, and not merely to have a life happen to one, depends on being able to express one's autonomy and on being protected against persons who are unprepared to mitigate their action in the light of the interests of others. Of course, any plausible theory of what is valuable to a person would include the ability to act on the basis of one's preferences and desires. But that is because autonomous action is valuable to persons understood as planning agents who bear a special relationship of ownership and responsibility to how their life goes, and not because people have a taste for autonomy.

In the context of this brief Review, I cannot develop this line of argument in detail. I have said a bit more about it elsewhere, and views of this sort are familiar in the literature more broadly. The idea here is that once we acknowledge that human welfare matters because of something about what it is to be a person—that is, to be an agent capable of living a life of one's own, where how one's life goes is in part a matter of what one does and not just a function of what happens to one—it is obvious that the very same kinds of considerations that explain the value of welfare explain our strong deontic intuitions as well. Principles that restrict the extent to which we can pursue our own interests without regard for the impact of our
actions on the interests and rights of others express a commitment to this ideal of the person. We can think of distributive justice in roughly the same way. It is a precondition of one's life being something one does, rather than something that happens to one, that one have resources at one's disposal adequate to that end. One who is completely the victim of misfortune and bad luck is robbed of the capacity to realize oneself or one's personality in the world. Welfare matters because the self-respect and dignity of persons matter. The conditions of self-respect and human dignity require us to mitigate our actions in ways that take into account the interests of others, and to regulate our conduct by norms that fairly and justly adjudicate among those competing interests.

Once we realize that welfare is connected to a person's interests—what is good for him, and not merely to what he desires or to his gratification or joy—it should be clear that whatever it is in that account that explains the value of welfare explains as well the importance of the law's regulating human affairs according to various principles of justice and fairness. It is something about people, and not something about realizing desires or gratifying psychological states, that makes human welfare valuable. But whatever it is about persons that ultimately warrants concern for human welfare warrants the view that justice must regulate affairs between persons. It is not that justice is a constituent of welfare or welfare a constituent of justice. Rather, both are important and distinct reflections of the dignity and importance of persons. Any theory of the law that would direct us to evaluate our practices by considering only welfare or justice and not the other could do so only by impoverishing the idea of the person. In doing so, it would indict itself more than any critic, sympathetic or otherwise, could.