

PROLOGUE

In this book, we ask what criteria ought to guide social decisionmaking. Our thesis is that social decisions should be based *exclusively* on their effects on the welfare of individuals—and, accordingly, should not depend on notions of fairness, justice, or cognate concepts.

This thesis runs counter to conventional wisdom and much academic evaluation of social policy, conflicts directly with the views of most twentieth-century moral philosophers, and may well grate against the reader's intuition. For example, most individuals would accord at least some weight to the idea that punishment should fit the crime, even when the fitting punishment is not the most effective way to reduce criminal activity. And most people would favor providing a level playing field for competitors in the marketplace because fairness to participants demands it, even when it might not be most conducive to productivity.

Let us describe how we—two economists (one also a lawyer) on a law school faculty—came to develop our thesis. Our training as economists as well as our natural inclinations lead us to evaluate policies based on what ultimately matters to people, that is, on how policies affect individuals' well-being. But we have encountered particularly acute criticism of this approach in the legal academy, where invocations of fairness and justice are the norm. As a consequence, we have been thinking about and responding to the criticism since we joined the law faculty in the early 1980s. In 1995, we decided to embark on a systematic investigation of the conflict between fairness-based analysis and that grounded in an exclusive concern for individuals' well-being.

Although our initial focus in our work was on normative legal analysis, it was apparent to us from the outset that our subject is not limited to the

law but rather concerns the most general questions of social policymaking and moral philosophy. Moreover, in the basic legal settings that we examine, the relevant notions of fairness are considered mainly by moral philosophers. In fact, the settings that we investigate span a number of the most basic subjects addressed in moral philosophy: What is *A*'s obligation to *B* when *A* wrongfully injures *B*? When are individuals obligated to keep promises, and what duty do promisors owe to promisees when they do not? What protections should individuals receive when the state's apparatus is used to resolve disputes? When and to what extent should the state impose punishment? These questions have engaged philosophers from Aristotle to Hume and Kant, to Ross and Hare and Rawls.

In light of the foregoing, we were led to broaden and deepen our normative analysis. What we have produced, although it is framed in legal settings, is an expansive inquiry into the principles that should guide public policy. The reader might naturally wonder what makes our analysis of this subject distinctive. We believe that our book makes three types of contributions.

First, we offer a number of arguments indicating that the conflict between notions of fairness and welfare is much sharper than has been appreciated. One of our conclusions in this regard is particularly striking: Under any method of evaluating social policy that accords positive weight to a notion of fairness, there must exist situations in which *all* individuals will be made worse off. To illustrate, consider a principle of fairness under which victims have the right to obtain compensation from their injurers. Suppose that all individuals are identically situated in the sense that each is just as likely to gain as a victim (receiving compensation) as he is to lose as an injurer (paying compensation). Now, viewing the matter prospectively, before individuals know whether they will turn out to be victims or injurers, it is clearly possible that upholding the right to compensation will make everyone worse off if implementing the compensation requirement is costly. In the early stages of our research, we constructed a variety of such examples showing that endorsement of particular notions of fairness will sometimes make everyone worse off. These results, which we found somewhat surprising, led us to inquire about their generality. We subsequently developed two formal arguments, presented in this book as well as in separate articles, establishing that virtually any method of evaluation that gives weight to notions of fairness will sometimes lead to choices that make all persons worse off. (The foregoing argument and others that we offer do not apply to many principles concerning the distribution of income or well-being, which are often advanced under the rubric of fairness and justice, because they are in fact embraced by a welfare-based approach.)

This conclusion has strong implications. Most individuals—including

many of the philosophers we have queried—would not readily endorse a principle of fairness if doing so implies (as it does) that it may be deemed socially good to make everyone worse off. It is, after all, difficult to understand the point of a notion of fairness if every person to whom one presumably seeks to be fair may be made worse off as a result. For the many who ground moral principles in individuals' freedom and autonomy, the foregoing is especially troublesome because it is hard to argue that all individuals, if given the opportunity, would freely choose to make themselves worse off. Furthermore, as we explain, one of the arguments underlying our conclusion poses an important challenge to those who adhere to the Golden Rule, Kant's categorical imperative, or the requirement that ethical principles be acceptable to parties situated behind a veil of ignorance.

Second, we undertake a thorough and systematic analysis of a number of notions of fairness in a variety of important, paradigmatic contexts, including, as noted, those in which one person wrongfully harms another, in which persons promise to perform contracts, in which procedures are provided for legal redress, and in which the state imposes punishment. By focusing on such basic cases, by making our analysis concrete, and by examining in detail the conditions under which implications of fairness and welfare are aligned and when they conflict, we believe that we are able to achieve a significantly better understanding of fairness-based evaluation than has been presented elsewhere. In particular, we consider at length whether the commonly proffered rationales for the notions, or others that we might imagine, can withstand scrutiny when the full ramifications of fairness principles are made clear. An additional benefit of our investigation is that very different notions of fairness are explored in parallel. Because most of our arguments apply to each of them, it appears that we have identified generic deficiencies in notions of fairness rather than merely *ad hoc* shortcomings of this or that notion.

Third, we consider a number of ways in which the broad appeal that many notions of fairness possess can be reconciled with our thesis that they should not be accorded independent weight in the evaluation of social policy. Most importantly, we examine the nature and origins of our moral instincts and intuitions, which provide the implicit or explicit grounding for many arguments of twentieth-century philosophers. This discussion, it should be emphasized, is descriptive and scientific in nature, not a matter of ethical theory. The subject of the origin of our moral feelings, though largely ignored in modern philosophical literature, has received substantial attention in the past from such prominent philosophers as Hume, Mill, and Sidgwick, and was explored by Darwin as well. Drawing upon these thinkers as well as modern scholars in social, cognitive, and natural sciences, we offer a syn-

thetic, welfare-based account of the origin of notions of fairness. Specifically, we suggest that these notions, which may be inculcated or innate, have emerged because of their functional value in governing our behavior in everyday social interactions; individuals' belief in notions of fairness reduces their inclination to engage in harmful behavior and otherwise leads them to act in ways that promote human welfare.

To the extent that the welfare-based explanation for our moral instincts and intuitions is valid, it makes no sense—indeed, it would be paradoxical—to impute significance to the notions of fairness that embody them when doing so comes at the expense of human welfare. Moreover, there is an often-overlooked distinction in contexts that helps to explain why seemingly appealing notions of fairness would sometimes lead us astray in designing social policy: The underlying moral intuitions and instincts arose for the purpose of regulating behavior in informal interactions in everyday life, whereas the realm of policymaking often differs from that of everyday life in important ways that we delineate.

We hope that this book will be of interest to a wide spectrum of readers. Because our exploration of notions of fairness and welfare addresses the normative foundation of public policy, it should be relevant to the work of social scientists, policy analysts, and others who are interested in the criteria that should govern social decisions. The book should be of particular interest to moral philosophers because we take issue with many of them in a fundamental way. We have attempted to consider their writing in depth, as our notes and list of references should make evident, and we relate our thesis to a wide range of basic topics in moral philosophy, including the relevance of two-level moral theories, the doctrine of double effect, the meaning of well-being, the problem of objectionable preferences, and the permissibility of ever punishing the innocent.