

BEYOND EFFICIENCY AND PROCEDURE: A WELFARIST THEORY OF REGULATION

MATTHEW D. ADLER*

I.	INTRODUCTION.....	241
II.	NEOCLASSICAL THEORIES OF REGULATION.....	244
	A. <i>Characterizing Neoclassicism</i>	244
	B. <i>Kaldor-Hicks Efficiency: Does It Have Moral Significance?</i>	248
	1. <i>Putative Defenses of Kaldor-Hicks Efficiency</i>	249
	2. <i>Some Spurious Criticisms of Kaldor-Hicks Efficiency</i>	259
	C. <i>The Preference-based View of Well-being</i>	262
	1. <i>Problems with the Preference-based View</i>	262
	2. <i>Alternative Theories of Well-Being</i>	264
III.	PROCEDURALIST THEORIES OF REGULATION.....	267
	A. <i>Characterizing Proceduralism</i>	267
	B. <i>The Defense Based on Moral Skepticism</i>	269
	C. <i>The Foundationalist Defense</i>	272
	D. <i>The Conventionalist Defense</i>	278
	E. <i>The Expressivist Defense</i>	281
	F. <i>The Interest-Based Defense</i>	283
IV.	A WELFARIST THEORY OF REGULATION.....	288
	A. <i>Interpersonal Comparisons and the Determinacy of Overall Well-Being</i> ...	289
	B. <i>The Moral Force of Overall Well-Being</i>	302
	C. <i>Deontology, Perfections, and Distribution: A Note on Nonwelfarist Considerations</i>	313
	D. <i>Welfarism and Neoclassicism: Some Differences</i>	319
	1. <i>Efficient Welfare Losses and Inefficient Welfare Gains</i>	320
	2. <i>Restricted and Unrestricted Preferences</i>	324
	3. <i>Incomparability and Precise Equality</i>	328
	4. <i>The Possibility of Paternalism?</i>	330
	E. <i>Welfarism and Proceduralism: A Large Difference</i>	332
V.	CONCLUSION.....	336

I. INTRODUCTION

What justifies regulatory policy? What are the moral criteria by reference to which regulatory choices and institutions are properly assessed? These questions have yet to be convincingly answered, or so I will argue in this Article. My focus will be legal scholarship, since it is legal scholars who, in recent years, have paid the most sustained attention to the problem of justifying regulation—the problem of generating a moral theory of regulation in light of which general regulatory approaches, specific regulatory decisions, the design of

* Professor of Law, University of Pennsylvania Law School. Many thanks to Rob Atkinson, Brian Bix, Howard Chang, Eric Posner, and Dan Rodriguez, to participants in the Florida State University Law School Symposium on Regulatory Theory and the University of Pennsylvania Law School faculty retreat, and to students in Penn’s Topics in Legal Theory course, for their extremely helpful comments. Many thanks, too, to Mark Seidenfeld and Heidi Hurd for giving me an opportunity to present drafts of this Article at Florida State and Penn.

regulatory agencies, and all other aspects of regulation can be evaluated as good or bad, right or wrong.

I will say nothing here about positive regulatory theories—theories that attempt to predict what regulatory outcomes will be, given (more or less) simple assumptions about human motivation, human beliefs, and so on—except to make the obvious point that no positive theory can tell us whether the outcomes it predicts are to be embraced or avoided. For example, a positive theory may reliably predict that a regulatory agency with jurisdiction over a concentrated industry will choose outcomes which inefficiently benefit the industry, at the expense of diffuse and unorganized groups such as consumers or taxpayers. But this prediction is not yet grounds for dismantling or constraining the agency. It is not yet grounds for doing *anything*.

Imagine that someone in a position of power over the agency (a legislator, an administrator, or the President, for example) has the capability to perform an action such that, if the person performs that action, the agency will produce an efficient but otherwise morally neutral outcome, relative to the outcome that the agency will produce if the powerful actor stays her hand. Thus our positive theory predicts. Then the actor ought to take the action if, and only if, efficiency itself has moral import. If it does not—if the efficient cast of an outcome is not (wholly or partly) constitutive of its goodness or badness, rightness or wrongness—then the action will be a matter of indifference. For if efficiency is *not* (wholly or partly) constitutive of goodness or badness, rightness or wrongness, then an actor has no more reason to choose one of two outcomes that differ only with respect to efficiency than she has to choose one of two outcomes that differ only with respect to other morally trivial properties. The fact that one outcome (O_1) is more efficient than another (O_2) would no more justify the powerful actor in producing O_1 than the fact that in O_1 a particular blade of grass is located at a particular spatio-temporal location, and that in O_2 it is located somewhere else. But whether efficiency is (wholly or partly) constitutive of goodness or badness, rightness or wrongness, is a question of moral theory, not positive theory.

Does efficiency indeed possess bedrock moral significance? The reader may be alarmed at my use of the adjective “bedrock.” If so, she should reconsider her alarm. “Bedrock” adds nothing, besides emphasis, to “moral significance.” A moral theory purports to identify the features of actions or outcomes that possess moral significance or, equivalently, bedrock moral significance. Where an actor can choose one action (A_1) leading to O_1 or another action (A_2) leading to O_2 , a moral theory purports to isolate the features of A_1 and A_2 , O_1 and O_2 , such that whether the actor is, all things considered, justified

in choosing A_1 rather than A_2 is a collective function of those features. If the features are thus isolated by the theory, then they have bedrock significance (within that theory); if not, then they do not.

In this Article I will criticize the two types of moral theories of regulation that have been most influential within American legal scholarship. One type, which I will call "neoclassical," attributes bedrock significance to efficiency. The second type, which I will call "proceduralist," does not do so, but instead attributes bedrock significance to the procedures by which regulatory agencies reach their decisions. The neoclassical type of theory is epitomized by Stephen Breyer's well-known book, *Regulation and Its Reform*,¹ and is reflected in the leading casebooks on regulation.² The proceduralist type of theory includes the "interest representation" theory delineated (if not defended) by Richard Stewart in his seminal article *The Reformation of American Administrative Law*;³ the "civic republican" theories defended by Mark Seidenfeld,⁴ Cass Sunstein,⁵ and others;⁶ and, most recently, the theory of collaborative governance advanced by Jody Freeman.⁷

Neoclassical and proceduralist theories are both flawed; but, as I will suggest, it is proceduralist theories, articulated largely in response to neoclassicism, that are the most deeply flawed. Efficiency does not possess moral significance, but *overall well-being* does, and the two criteria are significantly coextensive. By contrast, governmental procedures, particularly the procedures of regulatory agencies, have no normative significance. Two governmental decisions which differ only in the procedures by which they were reached are

1. STEPHEN BREYER, *REGULATION AND ITS REFORM* (1982).

2. See, e.g., JEFFREY L. HARRISON ET AL., *REGULATION AND DEREGULATION: CASES AND MATERIALS* (1997); RICHARD J. PIERCE, JR., *ECONOMIC REGULATION: CASES AND MATERIALS* (1994); LOUIS B. SCHWARTZ ET AL., *FREE ENTERPRISE AND ECONOMIC ORGANIZATION: GOVERNMENT REGULATION* (6th ed. 1985); SIDNEY A. SHAPIRO & JOSEPH P. TOMAIN, *REGULATORY LAW AND POLICY: CASES AND MATERIALS* (2d ed. 1998).

3. Richard B. Stewart, *The Reformation of American Administrative Law*, 88 HARV. L. REV. 1667 (1975).

4. See Mark Seidenfeld, *A Civic Republican Justification for the Bureaucratic State*, 105 HARV. L. REV. 1511 (1992).

5. See CASS R. SUNSTEIN, *AFTER THE RIGHTS REVOLUTION: RECONCEIVING THE REGULATORY STATE* (1990); Cass R. Sunstein, *Beyond the Republican Revival*, 97 YALE L.J. 1539 (1988) [hereinafter Sunstein, *Republican Revival*]; Cass R. Sunstein, *Factions, Self Interest, and the APA: Four Lessons Since 1946*, 72 VA. L. REV. 271 (1986); Cass R. Sunstein, *Interest Groups in American Public Law*, 38 STAN. L. REV. 29 (1985).

6. See Steven G. Gey, *The Unfortunate Revival of Civic Republicanism*, 141 U. PA. L. REV. 801 (1993) (citing sources); see also Steven P. Croley, *Theories of Regulation: Incorporating the Administrative Process*, 98 COLUM. L. REV. 1, 76-86 (1998) (same); Note, *Civic Republican Administrative Theory: Bureaucrats as Deliberative Democrats*, 107 HARV. L. REV. 1401 (1994) (same).

7. See Jody Freeman, *Collaborative Governance in the Administrative State*, 45 UCLA L. REV. 1 (1997).

morally identical. Proceduralism, because it says otherwise, is fundamentally misdirected.

Part II criticizes neoclassical theories. Part III criticizes proceduralist theories. Part IV outlines and defends an alternative type of theory—a welfarist theory of regulation. A welfarist theory gives moral significance (if not conclusive significance) to the criterion of overall well-being. Welfarism and neoclassicism may not differ too much in their practical implications, but neither are they practically indistinguishable. In Part IV I describe some of the noteworthy ways in which the recommendations of welfarism and neoclassicism differ.

Welfarism is, in effect, refurbished neoclassicism. Efficiency is an untenable moral foundation for the practice of regulation, but that foundation can be replaced without too much shifting of the structure above. Proceduralists and others who properly criticize neoclassicism go astray in thinking that regulatory agencies should make a large-scale, rather than an incremental, shift away from the pursuit of economic efficiency.

II. NEOCLASSICAL THEORIES OF REGULATION

A. Characterizing Neoclassicism

By a “neoclassical” theory of regulation, I mean roughly a theory grounded in the normative premises of modern welfare economics. More precisely, it is a theory with the following two elements: (1) the adoption of Kaldor-Hicks efficiency as one, if not the only, moral criterion by which to evaluate regulatory choices, policies, and institutions;⁸ and (2) the adoption of the preference-based account of well-being, which holds that whether a particular person is made better or worse off by a choice, policy, or institution wholly depends upon the extent to which that person’s preferences are satisfied.⁹ Neoclassical theories are espoused by such classics as Breyer’s *Regulation and Its Reform*¹⁰ and Alfred Kahn’s *The Economics of Regulation*,¹¹ by most of the leading legal casebooks on regulation;¹² by much law-and-

8. See, e.g., J.R. Hicks, *The Foundations of Welfare Economics*, 49 *ECON. J.* 696 (1939); Harold Hotelling, *The General Welfare in Relation to Problems of Taxation and of Railway and Utility Rates*, 6 *ECONOMETRICA* 242 (1938); Nicholas Kaldor, *Welfare Propositions of Economics and Interpersonal Comparisons of Utility*, 49 *ECON. J.* 549 (1939).

9. See, e.g., DANIEL M. HAUSMAN & MICHAEL S. MCPHERSON, *ECONOMIC ANALYSIS AND MORAL PHILOSOPHY* 71-83 (1996) (describing and criticizing the preference-based view of well-being held by welfare economics); L.W. SUMNER, *WELFARE, HAPPINESS, AND ETHICS* 113-22 (1996) (same).

10. BREYER, *supra* note 1.

11. 1 ALFRED E. KAHN, *THE ECONOMICS OF REGULATION: PRINCIPLES AND INSTITUTIONS* (1988).

12. See sources cited *supra* note 2. A similar approach is adopted by one of the leading contemporary textbooks on policy analysis. See EDWARD M. GRAMLICH, *A GUIDE TO BENEFIT-COST ANALYSIS* 9-29 (2d ed. 1990).

economics scholarship on regulation;¹³ and by Anthony Ogus's book on regulation,¹⁴ which is the most recent attempt by a legal scholar to provide something approaching a comprehensive, normative theory of regulation. Note that the term "neoclassical," as I have defined it, denotes a class of normative theories, rather than a unique theory. Two neoclassicists who agree as to the moral significance of efficiency, and as to the equivalence of welfare and preference satisfaction, may disagree about whether some further moral criterion (besides efficiency) is also relevant in evaluating regulatory choices, options, and institutions.¹⁵ One neoclassicist may think that the environment has intrinsic value, apart from human welfare; another may deny that. Or, one neoclassicist may think that the fair distribution of welfare, as well as Kaldor-Hicks efficiency, is an important thing; another neoclassicist may be solely concerned with efficiency.¹⁶ It is certainly not the case that Breyer, Kahn, Ogus, and all the scholars I have just styled as neoclassicists adopt the same, unique theory of regulation; but it is true, I suggest, that the otherwise divergent theories they propose share a common commitment to efficiency and to preferences.

This is not immediately obvious. It is a matter of interpretation, or at least sometimes it is—since the neoclassicist's commitment to efficiency and to preferences is sometimes implicit or even obscured rather than laid plain to view.¹⁷ Nonetheless, this commitment is plausibly inferred from other claims that Breyer, Ogus, and the other neoclassicists do make quite explicitly and centrally. Consider the "market failure" framework that these scholars adopt, within which monopolization and spillovers (externalities) are seen to be two para-

13. See, e.g., A. MITCHELL POLINSKY, AN INTRODUCTION TO LAW AND ECONOMICS 7-10 (2d ed. 1989) (stating that law and economics focuses on the "efficiency" of legal rules, with efficiency defined in the Kaldor-Hicks sense as "the relationship between the aggregate benefits of a situation and the aggregate costs of the situation"); Chris William Sanchirico, The New Efficiency Rationale: An Internal Critique (1999) (unpublished manuscript, on file with author) ("For as long as Law and Economics has played a formative force in legal scholarship, its principal mode of analysis has been to evaluate legal rules according to the sole criterion of 'efficiency': that is, to judge rules by the simple sum of the costs and benefits that they impose on the population without regard to how those costs and benefits are distributed among different individuals."), available at <http://www.cstone.net/~csanchir/Taxes2.pdf> (visited Oct. 3, 2000).

14. ANTHONY I. OGUS, REGULATION: LEGAL FORM AND ECONOMIC THEORY (1994).

15. See BREYER, *supra* note 1, at 32-33 (citing various possible noneconomic rationales for regulation); HARRISON ET AL., *supra* note 2, at 30 (same); KAHN, *supra* note 11, at 189-93 (same); OGUS, *supra* note 14, at 46-54 (same); PIERCE, *supra* note 2, at 11-50 (same); SCHWARTZ ET AL., *supra* note 2, at 79 (same); SHAPIRO & TOMAIN, *supra* note 2, at 313-55, 557-657 (same).

16. See *infra* Part IV.C (discussing the arguable moral relevance of criteria other than overall well-being, including perfectionist and distributive criteria).

17. But not always. Some neoclassicists explicitly adopt efficiency, in the Kaldor-Hicks sense, as a criterion for assessing regulatory policy. See, e.g., HARRISON ET AL., *supra* note 2, at 20-25; OGUS, *supra* note 14, at 29.

digmatic failures that justify some kind of regulatory response.¹⁸ Anyone who thinks that the state is *prima facie* justified in moving a market from the price-output point reached under monopoly (P_m), to the price-output point reached under competitive conditions (P_c), is plausibly interpreted as adopting a commitment to Kaldor-Hicks efficiency. Why? Because the move from P_m to P_c is necessarily Kaldor-Hicks efficient,¹⁹ but there is no other plausible moral criterion that necessarily counts the move from P_m to P_c as a good thing.²⁰ That move is not Pareto-efficient,²¹ since the monopolist is made worse off by the move. It does not necessarily increase overall well-being, since money has (or can have) differential welfare productivity across persons. Although the monopolist's foregone profits at P_c are less, in dollar terms, than the consumers' gained surplus, the consumers may be wealthier than the monopolist, or otherwise less "productive" at converting dollars into welfare, so that P_m is actually welfare maximizing relative to P_c . Even more obviously, the move from P_m to P_c is not necessarily better in light of distributive criteria (egalitarian criteria, for example) since the monopolist may again be poorer in overall wealth or welfare than the consumers, or otherwise more deserving in light of plausible standards for fair distribution. Nor, clearly, is it necessarily better in light of deontological criteria, since some anti-monopoly policies could move society from P_m to P_c but also violate the monopolist's deontological rights, or in light of perfectionist criteria, since they have no connection at all to human well-being. The consumers at P_c have enjoyed a welfare gain that is large enough, in dollar terms, to compensate for the monopolist's welfare loss. In short, it is Kaldor-Hicks efficient, and that alone explains why the state might be *prima facie* justified in choosing P_c .

18. See BREYER, *supra* note 1, at 15-16, 23; HARRISON ET AL., *supra* note 2, at 124-65, 191-94; OGUS, *supra* note 14, at 30-33, 35-38; PIERCE, *supra* note 2, at 12-21, 24-32; SHAPIRO & TOMAIN, *supra* note 2, at 32-34; *cf.* KAHN, *supra* note 11 (focusing on monopoly); SCHWARTZ, *supra* note 2, at 63-65 (same).

19. More precisely, the move is necessarily Kaldor-Hicks efficient if the welfare-relevant preferences are taken to be those of consumers and producers, see *infra* Section IV.D.2, and if special cases such as "natural monopoly" are placed to one side.

20. To quote Richard Posner:

[W]hen economists say that monopoly is inefficient, they mean inefficient in the Kaldor-Hicks or wealth-maximization, not the Pareto, sense. . . . The loss [in monopoly] is clear from a wealth-maximization standpoint: the sum of consumer and producer surplus is less under monopoly than under competition. . . . Thus a move from monopoly to competition would satisfy the Kaldor-Hicks or wealth-maximization criterion for a gain in efficiency. But it would not satisfy the criterion of Pareto superiority, because the monopolist would be worse off.

RICHARD A. POSNER, THE ECONOMICS OF JUSTICE 91-92 (1981).

21. One outcome is "Pareto-efficient" or "Pareto-superior," relative to a second, if at least one person is better off with the first outcome and no one is worse off. In that case, the second outcome is "Pareto-inefficient" or "Pareto-inferior" relative to the first. If one outcome is neither Pareto-inferior nor Pareto-superior to another, the two are "Pareto-noncomparable."

A similar point can be made about the identification of externalities (spillovers) as a key scenario warranting regulatory intervention. Neoclassicists see externalities as troubling because externalities can be Pareto-inferior. If the external costs of an activity are large enough, the state of affairs in which the actor performs the activity will be Pareto-inferior to some other state in which the actor refrains from the activity. Indeed, the link between externalities and Pareto-inferiority is inherent in the description of externalities as a “market failure”; for what is special about an ideally functioning market just is the fact that its outcomes are Pareto-optimal.²² By contrast, it is clear that an ideally functioning market—one in which (among other things) actors are perfectly informed, conditions are competitive, and all goods are private—may still fail to maximize well-being, achieve distributive justice, comply with deontological requirements, or advance perfectionist goals.

But neoclassicists do *not* suggest that an activity which is Pareto-inferior by virtue of its external costs should be regulated only if the regulator arranges or anticipates a side-payment from the cost-bearers to the actor. Rather, they think (or seem to think) that a Pareto-inferior activity is justifiably regulated without more—without the actual payment of compensation to the actor. And this view implies a commitment to Kaldor-Hicks efficiency, just like the view that a monopolized market is justifiably moved from P_m to P_c even if the monopolist is not actually compensated for the move.

What about the second component of neoclassicism, the adoption of a preference-based account of well-being? The preference-based view of well-being is even more deeply entrenched in welfare economics than the criterion of Kaldor-Hicks efficiency,²³ and often neoclassicists espouse that view quite plainly.²⁴ Where they do not, the view can plausibly be inferred from other claims.

First, it can be inferred from their adoption of “imperfect information” rather than, say, “objective badness” or “failure of pleasure,” as a prime scenario where regulation may be justified.²⁵ Consider the case where a purchaser (P) is about to buy a good or service (G). If the preference-based view of well-being obtains, then regulators justifiably intervene here, in light of P 's well-being, only if (1) P is imperfectly informed about the consequences of buying G , or (2) P is ir-

22. An outcome is Pareto-optimal if there are no other outcomes relative to which it is Pareto-inefficient.

23. See sources cited *infra* note 29 (citing theoretical and applied welfare economists who reject Kaldor-Hicks criterion, but who do not generally reject the preference-based account of well-being).

24. See, e.g., OGUS, *supra* note 14, at 38.

25. See BREYER, *supra* note 1, at 26-28; HARRISON ET AL., *supra* note 2, at 210-19; OGUS, *supra* note 14, at 38-41; PIERCE, *supra* note 2, at 32-33; SHAPIRO & TOMAIN, *supra* note 2, at 34-35.

rational. By contrast, the regulatory theorist's focus on "imperfect information" would be quite puzzling if the theorist held an objective-list view of well-being (such that *P*'s well-being consists in his achievement of one or more objective goods, quite apart from his preferring or not preferring that achievement) or a hedonic view of well-being (such that *P*'s well-being consists in his experiencing one or more types of pleasurable mental states, quite apart from his preferring or not preferring that experience). Perfectly informed and perfectly rational persons can still fail to realize objective goods or experience an optimal mix of pleasures, insofar as they just do not care about the goods or the pleasures.²⁶

Second, the neoclassicist's commitment to a preference-based view of well-being is plausibly inferred from his distaste for paternalism. Some kind of antipaternalism, ranging from mild skepticism to fierce hostility, is characteristic of virtually all of the scholarship under consideration here.²⁷ Note that antipaternalism is one immediate implication of a preference-based view of well-being. On that view, if *P* prefers the state of affairs which results from his buying *G* to the state which results from his not buying *G*, then government's paternalistic intervention to prevent *P* from buying *G* cannot make him better off. By contrast, on an objective-list or a hedonic view of welfare, governmental intervention in the teeth of *P*'s preferences *can* still be welfare improving for *P*. A certain kind of antipaternalism may follow from objectivism—if the objectivist counts autonomy as one of the goods constitutive of welfare—but the link to antipaternalism is less fundamental here than within a preference-based view.

B. Kaldor-Hicks Efficiency: Does It Have Moral Significance?

So much for my interpretive claim that Breyer, Ogus, and the other neoclassicists are committed to (1) Kaldor-Hicks efficiency as a moral criterion (if not the only criterion) by which to evaluate regulatory choices, policies, and institutions and (2) a preference-based account of well-being. What is wrong with these commitments? Let me

26. It might be objected that objective goods are those goods that all persons, under conditions of perfect information and rationality, would prefer—and thus that perfectly informed and rational persons *would* realize objective goods. See *infra* text accompanying note 151. I would respond by noting, first, that objective goods are those things that perfectly informed and rational persons would *restrictedly* prefer, i.e., prefer for themselves; a perfectly informed and rational person could act contrary to her own welfare and deny herself objective goods. Second, even in the case where preferences are "restricted"—where action is self-interested—a perfectly informed and rational person will have a *prima facie* preference to realize objective goods, but may prefer, all things considered, to realize her own, idiosyncratic conception of welfare. See *infra* Section II.C.

27. See, e.g., BREYER, *supra* note 1, at 33-34; OGUS, *supra* note 14, at 51-53. *But see* SHAPIRO & TOMAIN, *supra* note 2, at 35.

start with efficiency, since I think the commitment to efficiency is more deeply mistaken than the commitment to preferentialism. There is (as we shall see) a large grain of truth in preferentialism, but the thought that Kaldor-Hicks has normative import is plain wrong.²⁸ These days, theoretical welfare economists generally *reject* Kaldor-Hicks efficiency as a standard for ranking social states.²⁹ And the considerations against efficiency as a moral criterion were fully articulated in the legal literature almost twenty years ago, in contributions by Ronald Dworkin,³⁰ Jules Coleman,³¹ Mark Kelman,³² and others.³³ Nonetheless, the efficiency criterion lingers on in the works of neoclassical regulatory theorists, and more generally in much of law and economics. It thus bears repeating just why Kaldor-Hicks efficiency should be removed from our normative vocabulary. It lacks the basic status of Pareto-efficiency, overall well-being, distributive considerations, deontological requirements, and (perhaps) perfectionist value.

1. Putative Defenses of Kaldor-Hicks Efficiency

Consider the following putative defenses of Kaldor-Hicks efficiency, which at one time or another have been presented by economists or legal scholars. First, Kaldor-Hicks efficiency is sometimes linked to overall well-being.³⁴ But it is clear that the two can di-

28. For a more compact argument defending the related point that cost-benefit analysis lacks normative import, see Matthew D. Adler & Eric A. Posner, *Rethinking Cost-Benefit Analysis*, 109 YALE L.J. 165, 187-94 (1999).

29. See, e.g., ROBIN BOADWAY & NEIL BRUCE, WELFARE ECONOMICS 96-101 (1984); Charles Blackorby & David Donaldson, *A Review Article: The Case Against the Use of the Sum of Compensating Variations in Cost-Benefit Analysis*, 23 CANADIAN J. ECON. 471 (1990); Stephen Coate, *Welfare Economics and the Theory of Project Evaluation* (April 1996) (unpublished manuscript on file with author). For a frank rejection of the Kaldor-Hicks criterion by a leading textbook in *applied* welfare economics, see COST-BENEFIT ANALYSIS 6 (Richard Layard & Stephen Glaister eds., 2d ed. 1994), which states, "[T]here is no ethical justification for the Hicks-Kaldor criterion; where compensation will not be paid there seems no alternative to interpersonal comparisons of the value of each person's gains and losses."

30. See Ronald M. Dworkin, *Is Wealth a Value?*, 9 J. LEGAL STUD. 191 (1980).

31. See Jules L. Coleman, *Efficiency, Utility, and Wealth Maximization*, 8 HOFSTRA L. REV. 509 (1980).

32. See MARK KELMAN, A GUIDE TO CRITICAL LEGAL STUDIES 114-50 (1987).

33. See generally *Symposium on Efficiency as a Legal Concern*, 8 HOFSTRA L. REV. (1980); Lewis A. Kornhauser, *Wealth Maximization*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW (Peter Newman ed., 1998).

34. This seems to be the suggestion in J.R. Hicks, *The Valuation of the Social Income*, 7 ECONOMICA (n.s.) 105, 111 (1940), which states, "This [the Hicks criterion] would seem to be quite acceptable as a definition of increase in real social income." Anandarup Ray has also suggested a link to welfare:

Even though compensation tests are often claimed to be the foundation of traditional cost-benefit analysis, it is perhaps more sensible to interpret the traditional approach in a different way. If the gains and losses accruing to individuals are regarded as equal from the social point of view, then the [sum of com-

verge,³⁵ as I suggested above in my discussion of the regulation of monopoly. A simple illustration of the divergence between Kaldor-Hicks efficiency and overall well-being is the following. One so-called exchange economy (S_1) consists of a fixed stock of private goods ($G_1, G_2 \dots G_n$) plus an allocation of those goods to various persons. The total amounts of the goods are $G_1^*, G_2^* \dots G_n^*$. A second exchange economy (S_2) consists of the same goods ($G_1, G_2 \dots G_n$); however, the total amounts in this economy are $G_1^+, G_2^* \dots G_n^*$, and G_1^+ is bigger than G_1^* . In other words, the two economies have identical stocks of goods except that S_2 has more of G_1 .

Then S_2 is necessarily Kaldor-Hicks efficient relative to S_1 because, whatever the actual allocation in S_1 and S_2 , the goods in S_2 could always be reallocated (if reallocation were costless) to match the allocation in S_1 , leaving a surplus equaling the difference between G_1^+ and G_1^* . Surely, however, the move from S_1 to S_2 does not necessarily increase overall *well-being*. For example, assume that in both economies all of the goods $G_2 \dots G_n$ are allocated to a single individual Rich; in S_1 the stock of G_1 is divided equally among persons; but in S_2 all of good G_1 is allocated to Rich. Then a shift from S_1 to S_2 will decrease overall well-being, despite the acquired surplus, if it is the case that G_1 (like many goods) has declining marginal utility and the surplus (the difference between G_1^* and G_1^+) is not large enough to offset the declining marginal utility of G_1 . More generally, unless all of the goods have a constant (rather than declining or increasing) marginal utility that is the same amount for all persons, the shift from S_1 to S_2 will not be necessarily *welfare improving*, even though it is necessarily Kaldor-Hicks efficient.³⁶

pensating variations] measure would represent actual, rather than potential, changes in welfare.

ANANDARUP RAY, *COST-BENEFIT ANALYSIS: ISSUES AND METHODOLOGIES* 26 (1984); see also ROBERT SUGDEN & ALAN WILLIAMS, *THE PRINCIPLES OF PRACTICAL COST-BENEFIT ANALYSIS* 93-95 (1978) (suggesting that the Kaldor-Hicks criterion and distributive justice are different dimensions of "social welfare"); cf. Richard A. Posner, *Utilitarianism, Economics, and Legal Theory*, 8 J. LEGAL STUD. 103 (1979) (rejecting utilitarian defense of wealth maximization).

35. See David Copp, *The Justice and Rationale of Cost-Benefit Analysis*, 23 THEORY & DECISION 65, 74-77 (1987) (arguing that a project with net monetary benefits can decrease overall well-being); Peter Railton, *Benefit-Cost Analysis as a Source of Information About Welfare*, in VALUING HEALTH RISKS, COSTS, AND BENEFITS FOR ENVIRONMENTAL DECISION MAKING 55, 67-69 (P. Brett Hammond & Rob Coppock eds., 1990) (same).

36. The disjunction between Kaldor-Hicks efficiency and overall well-being is now widely recognized in the literature on welfare economics. For example, that disjunction is typically part of the case for the use of "distributive weights" in cost-benefit analysis, i.e., weighting factors that adjust an individual's monetary valuation of a governmental or regulatory project so as to reflect his marginal utility of money, or to reflect distributive considerations. See BOADWAY & BRUCE, *supra* note 29, at 271-91 (advocating distributive weighting); AJIT K. DASGUPTA & D.W. HEARCE, *COST-BENEFIT ANALYSIS: THEORY AND PRACTICE* 57-69 (1972) (same); D.W. HEARCE & C.A. NASH, *THE SOCIAL APPRAISAL OF PROJECTS: A TEXT IN COST-BENEFIT ANALYSIS* 27-38 (1981) (same); RICHARD O. ZERBE JR.

Second, Kaldor-Hicks efficiency is sometimes linked to Pareto-efficiency, via the claim that the surplus created by a particular Kaldor-Hicks efficient move can or will be redistributed through the tax system, making the move universally beneficial.³⁷ This claim is confused. Imagine that S_1 is a world where the regulator fails to initiate a particular regulatory policy, that S_2 is a world with the policy, and that S_2 is Kaldor-Hicks efficient relative to S_1 . S_2 has two variants, S_2^* and S_2^+ . In S_2^* , the taxing authorities engage in Paretian redistribution³⁸ of the surplus created by the move from S_1 to S_2 ; in S_2^+ they do not.³⁹ Why, now, is S_2 morally better than S_1 ? If the variant of S_2 that obtains is S_2^* , then we have a clear answer: S_2^* is Pareto-efficient relative to S_1 . But if the variant of S_2 that obtains is instead S_2^+ , then the question remains unanswered. The fact that S_2^* is morally better than S_1 by virtue of being Pareto-efficient hardly explains why S_2^+ , which is a different world than S_2^* , and which is not Pareto-efficient relative to S_1 , is also morally better than S_1 .

To put the point another way, the fact that a policy is Kaldor-Hicks efficient relative to the status quo merely entails that the policy *could* be transformed into a Pareto-efficient move via an ideal-

& DWIGHT D. DIVELY, BENEFIT-COST ANALYSIS: IN THEORY AND PRACTICE 236-53 (1994) (same); see also COST-BENEFIT ANALYSIS, *supra* note 29, at 6 ("Unless one dollar is equally valuable to all parties concerned, there are only two alternatives: to use some system of distributional weights or simply to show the net benefits to each party and let the policy maker apply his own evaluation.").

37. This seems to be the suggestion in Kaldor, *supra* note 8. See also E.J. MISHAN, COST-BENEFIT ANALYSIS 170-71 (4th ed. 1988) ("[T]here exist within Western societies economic institutions—such as progressive tax systems and welfare payments systems—which act over time to translate a potential into an actual Pareto improvement."); PEARCE & NASH, *supra* note 36, at 28 ("While ensuring [through application of the Kaldor-Hicks test] that we reach some position on the welfare frontier, we may use taxation and income supplementation as policy instruments to move along the frontier to our chosen distribution of income (citation omitted)."); SUGDEN & WILLIAMS, *supra* note 34, at 92 ("A project that satisfies the potential Pareto improvement criterion is one that the government can use, in conjunction with tax changes, to make everyone better off."). Louis Kaplow and Steven Shavell have also relied on the "surplus" rationale:

[M]oving to the efficient legal rule with an appropriate change in the income tax leaves all individuals equally well off but leaves the government with a surplus. With this additional revenue, the government can make each individual better off—for example, by lowering taxes by a fixed amount for each individual or by spending the funds on a public good that benefits everyone.

LOUIS KAPLOW & STEVEN SHAVELL, THE EFFICIENCY OF THE LEGAL SYSTEM VERSUS THE INCOME TAX IN REDISTRIBUTING INCOME 5 (National Bureau of Econ. Research Working Paper No. 4457, 1993).

38. By "Paretian" redistribution, I mean a redistribution from those who are better off in S_2 , as compared to S_1 , to those who would otherwise be worse off, such that the resulting outcome is in fact Pareto-superior to S_1 , i.e., everyone is better off.

39. There are a host of reasons, including corruption, administrative expense, democratic constraints, the absence of statutory authorization, and many others, that might prevent Paretian redistribution by the taxing authorities. The sheer fact that S_2 is Kaldor-Hicks efficient relative to S_1 does not guarantee that a Paretian redistribution will take place, since Kaldor-Hicks efficiency is defined in terms of hypothetical, not actual, redistribution from winners to losers.

ized, cost-free, lump-sum tax system. If a particular Kaldor-Hicks efficient policy will in fact be transformed into a Pareto-efficient move through the tax system (that is, we will move from S_1 to S_2^*), then the Kaldor-Hicks criterion is superfluous in explaining the moral attractiveness of the policy; that is more straightforwardly explained by the unimpeachable criterion of Pareto-efficiency. Conversely, if a particular Kaldor-Hicks efficient policy will *not* be transformed into a Pareto-efficient move (that is, we will move from S_1 to S_2^+), then something besides the Pareto criterion is needed to explain why the policy is a good thing.⁴⁰

Third, Kaldor-Hicks is sometimes defended through the notion of a *long-run* Pareto improvement.⁴¹ The idea, here, is not that the surplus from a particular Kaldor-Hicks move will be redistributed through the tax system, but that any particular Kaldor-Hicks move is part of a package of Kaldor-Hicks improvements such that this package is, collectively, better for everyone. To put the idea most crisply and plausibly: a general governmental (regulatory) practice of Kaldor-Hicks evaluation is Pareto-efficient relative to other practices.

There are a number of difficulties here. To begin with, the long-run or practice-level justification of Kaldor-Hicks does not constitute a full defense of Kaldor-Hicks as a moral criterion, since it fails to explain why one-off Kaldor-Hicks changes—changes that are efficient but occur in a world where no practice of Kaldor-Hicks evaluation exists, or where the existing practice does not reach such changes—are a good thing. More seriously, the claim that the practice of Kaldor-Hicks evaluation is necessarily or actually Pareto-efficient relative to other practices is simply untrue. One reason is

40. As Amartya Sen has explained:

There is a real motivational tension in the use of the logic of compensation for reading social welfare. If compensations are actually paid, then of course we do not need the compensation criterion, since the actual outcome already includes the paid compensations and can be judged without reference to compensation tests On the other hand, if compensations are not paid, it is not at all clear in what sense it can be said that this is a social improvement The compensation tests are either redundant or unconvincing.

Amartya Sen, *The Discipline of Cost-Benefit Analysis*, 29 J. LEGAL STUD. 931, 947 (2000); see also RAY, *supra* note 34, at 25 (“Unless decisionmakers can ensure that all project benefits are redistributed at no cost (in which case the Kaldor criterion is redundant), the link between potential welfare [in the sense of the Kaldor test] and actual welfare is not straightforward.”).

41. This idea is formalized in A. Mitchell Polinsky, *Probabilistic Compensation Criteria*, 86 Q.J. ECON. 407 (1972). E.J. Mishan also makes the “long-run” argument:

[E]ven though it may be the case that for each change sanctioned by the [potential] Pareto criterion a number of people will be made worse off, a succession of such changes is not likely to inflict losses on the same group. Over time, therefore, there can be a presumption that everyone or nearly everyone will be made better off by consistent application of the Pareto criterion.

MISHAN, *supra* note 37, at 171.

that particular persons may suffer large losses in particular rounds of Kaldor-Hicks evaluation, such that the gains they reap in other rounds are not substantial enough to make them net winners from the overall practice. Consider, for example, the hardly unrealistic case where a Kaldor-Hicks efficient decision (perhaps the decision of an environmental, health, or safety agency) causes grave physical harm or death to a few persons. Those persons could well be better off in a world where agencies follow a practice of Pareto-evaluation rather than Kaldor-Hicks evaluation and generally decline to issue a regulatory directive unless the outcome produced by that directive is truly Pareto-efficient relative to the status quo.⁴²

The attribution of Pareto-efficiency to the practice of Kaldor-Hicks evaluation is also false for a more subtle reason. There is a whole range of evaluative practices that governmental and regulatory agencies could follow other than Kaldor-Hicks evaluation and its salient alternatives (like Pareto-evaluation). Agencies could follow stranger alternatives such as approving projects only if they generate a certain kind of welfare benefit, or only if they benefit persons in a particular part of the country, and so on. But of course a practice of Kaldor-Hicks evaluation will not be Pareto-efficient relative to all alternatives. For example, if Joe Smith is a person, then the regulatory practice of Kaldor-Hicks evaluation will not be Pareto-efficient relative to the regulatory practice of making those and only those decisions that benefit Joe Smith, since Joe will be better off under this latter practice than under Kaldor-Hicks evaluation!

In short, there is a "Pareto-frontier" of regulatory practices along which practices are Pareto-noncomparable with each other, as well as an "interior" area of regulatory practices where each practice is Pareto-inferior to some practice on the frontier. Even if the practice of Kaldor-Hicks evaluation lies on the frontier, the criterion of Pareto-efficiency can at most explain why that practice should be chosen over certain "interior" practices, but not why it should be chosen over noncomparable practices (like the Joe Smith practice) that also lie on the frontier or that lie in the interior but are noncomparable with the Kaldor-Hicks practice. The long-run argument is not a defense of Kaldor-Hicks efficiency *simpliciter*? a defense against *all* alternative regulatory practices? but is at best a defense of that practice against a *particular* set of alternatives: namely, those prac-

42. This practice would not be well-defined unless the "status quo" outcome were defined, but presumably that could be done with reference to the action/inaction distinction (the status quo is the outcome in which the agency fails to act), or with reference to some common law baseline (the status quo is the outcome in which the set of legal rights, duties, and so on, is closer to the set defined by traditional property, contract, tort, and criminal law).

tices compared to which the Kaldor-Hicks practice is truly Pareto-superior.

Fourth, the Kaldor-Hicks criterion is sometimes linked to Pareto-efficiency by introducing risk or uncertainty.⁴³ The general idea seems to be that a Kaldor-Hicks change can prove (ex ante) Pareto-optimal once individuals or the government have probabilistic rather than perfect information. But a simple example will suggest that probabilistic considerations do not reliably support the Kaldor-Hicks criterion. Imagine a project that will affect ten people, five men and five women, producing some welfare benefit worth \$200 for each of the ten persons and some welfare setback worth \$300 for five of the ten persons. The defender of Kaldor-Hicks efficiency might argue that each of the ten persons calculates the expected net monetary impact of the project upon himself or herself as \$50, so that the project is, ex ante, a welfare improvement from the point of view of each person and therefore, ex ante, Pareto-efficient.⁴⁴

But this argument is spurious. First, although the project is necessarily Kaldor-Hicks efficient (since \$2000 is greater than \$1500), each person will not necessarily calculate his or her expected net gain from the project as \$50. *P₁* might have good grounds for the assessment that his risk of falling in the losing group is 90%. Similarly, if one person is a woman and it can be accurately predicted that four of the five women will be losers, then (knowing nothing more) the expected impact of the project upon her is -\$40 ((80% ? -100) + (20% ? 200)). The point here is that it takes much more than Kaldor-Hicks efficiency plus risk or uncertainty to ensure that the project is, ex ante, a welfare improvement from the point of view of each person; it takes a particular pattern of probability assessments as well. Absent that particular pattern, we are still left without any defense, ex ante or ex post, of the Kaldor-Hicks criterion.

43. See, e.g., Polinsky, *supra* note 41.

44. The argument would be that each person, not knowing who will suffer the setback, would estimate the impact of the project on her welfare thus:

Probability	Possible Outcome of Project	Weighted Value of Outcome
50%	-\$100	-\$50
50%	+\$200	+\$100
Total Expected Value of Project:		+50

In other words, she will consider she has a 50% chance of coming out \$200 ahead and a 50% chance of coming out \$100 behind. Thus the expected value of the project is \$50. If each person is risk-neutral, each should count this project as an ex ante welfare improvement.

Second, even if we put this point aside, the putative risk-based or uncertainty-based linkage from Kaldor-Hicks efficiency to Pareto-efficiency remains hazardous. Assume that each person in the example just described adopts the same estimate of his or her chances of benefit and setback from the project: 100% chance of a \$200 gross gain, 50% chance of a \$300 gross setback, and thus an expected net gain of \$50. So the project is, *ex ante*, welfare improving from the point of view of each person (at least on the simplifying assumption that each person is risk-neutral). This hardly implies that government should expect the project to be universally beneficial and, on those Paretian grounds, choose the project. After all, government knows for sure that there will be persons who are made worse off by the project—namely, the five unidentified persons who will in fact lose \$100 net. So the project is not, *ex ante*, an unequivocally good thing, from the point of view of government or of any other “observer” who takes into account the well-being of all ten persons. In short, the project is not, *ex ante*, Pareto-efficient; it would only be, *ex ante*, Pareto-efficient if government (incorrectly) expected that everyone would end up with a net gain from the project. Nor of course is it, *ex post*, Pareto-efficient, since in fact five persons end up losing. Thus the Pareto criterion provides no reason to choose the project.⁴⁵

To generalize: Even if the right pattern of probability assessments transforms a Kaldor-Hicks project into a project that is welfare-improving from the point of view of each affected person, such a project can *still* fail to be either *ex ante* or *ex post* Pareto-efficient. It can be clear *ex ante*, to a regulatory agency or some other government institution, that at least one person (not fully identified) will be harmed on balance by a government policy—that the policy will not be universally beneficial. This is true even if each person rationally expects to be a gainer and even if government has no further information *ex ante* than that available to each and all of the affected persons. For this important reason, among others, the putative link from Kaldor-Hicks to Pareto-efficiency, via risk or uncertainty, is a spurious one.⁴⁶

In the example just considered, the Kaldor-Hicks move at stake was a small-scale regulatory choice of one option over another, rather than a large-scale social choice of one regulatory practice over another, such as the practice of Kaldor-Hicks evaluation over the practice of Pareto or “Joe Smith” evaluation. However, my analysis ap-

45. This line of argument is advanced by John Broome, who claims that *ex ante* valuations of the risk of death are an inappropriate basis for cost-benefit analyses of governmental projects that cause death. See John Broome, *Trying to Value a Life*, 9 J. PUB. ECON. 91 (1978).

46. The defender of Kaldor-Hicks efficiency might use risk or uncertainty to draw a link from that criterion to the concept of *consent*, rather than to Pareto-superiority. This strategy is discussed and criticized below.

plies equally to the case of choosing practices. If one practice (P_1) is Kaldor-Hicks efficient relative to a second practice (P_2), and if information is probabilistic rather than perfect, it is still only contingently true that P_1 is, ex ante, welfare improving from the point of view of each person affected by the practices. And even if this *is* true, it may still be the case that P_1 is neither ex post nor ex ante a Pareto improvement. If at least one person loses from the practice, and if government has sufficient information to expect that, then P_1 is not universally beneficial either ex ante or ex post.

Fifth, the Kaldor-Hicks criterion is sometimes defended by reference to considerations of "consent."⁴⁷ The least plausible version of this defense runs as follows: since (1) the winners' gains from a Kaldor-Hicks move could be redistributed to the losers, producing a universal benefit, and since (2) everyone would consent to the Kaldor-Hicks move together with the Paretian redistribution, we can conclude that (3) everyone would consent to the Kaldor-Hicks move. Just because everyone would consent to the Kaldor-Hicks move plus Paretian redistribution (S'), and S' is related to the Kaldor-Hicks move alone (S) via a certain transformation (the redistribution), it hardly follows that everyone would or should or does consent to S .

A much better version of the consent-based defense of Kaldor-Hicks efficiency builds upon the idea, discussed above, that the introduction of risk or uncertainty can transform an individual's evaluation of a Kaldor-Hicks move. Consider the following argument for the Kaldor-Hicks criterion: (1) A Kaldor-Hicks move⁴⁸ is, ex ante, welfare-improving from the point of view of each affected person; (2) every person would consent to a move that is, ex ante, welfare-improving for her; and, therefore, (3) every person would consent to a Kaldor-Hicks move. One problem with this syllogism, already mentioned, is that a Kaldor-Hicks move may not be, ex ante, welfare-improving from the point of view of each affected person; that depends upon the pattern of probability assessments, and any a priori or conceptual link between Kaldor-Hicks and a particular pattern is spurious. But let us put that issue aside. Assume that, as it happens, all of the following are true of a particular Kaldor-Hicks move: (a) the move is not ex post a Pareto improvement, since at least one person will end up losing from it; (b) government does not believe ex ante that the move is a Pareto improvement, since (given its probability assessment) government believes that at least one person will lose

47. See, e.g., POSNER, *supra* note 20, at 88-115; Herman B. Leonard & Richard J. Zeckhauser, *Cost-Benefit Analysis Applied to Risks: Its Philosophy and Legitimacy*, in VALUES AT RISK 31, 34-36 (Douglas MacLean ed., 1986); Richard A. Posner, *The Ethical and Political Basis of the Efficiency Norm in Common Law Adjudication*, 8 HOFSTRA L. REV. 487, 491-95 (1980).

48. "Move" here could be a small-scale decision or a large-scale move from one regulatory practice to another.

from the project; and (c) each person, given her information, counts the move as an ex ante welfare improvement for her, and government knows that. The interesting question here is whether government has moral grounds to approve the move—grounds consisting in the fact that each person would consent to the move—even though it is neither ex ante nor ex post a Pareto improvement.

I think not, although the question is clearly not an easy one. If everyone affected actually consents to the move—if an actual contract is signed by all, or if the move is universally approved in a referendum—then there are moral grounds in favor of the project. *Actual* consent has moral force. But it is not clear that *hypothetical* consent does.⁴⁹ To begin with, it is clear that hypothetical consent does not have the *same* moral force as actual consent. For example, if I *would* or *should* contract with you to the effect that I do *A* if you do *B*, and you do *B*, then the grounds for forcing me to do *A* or pay damages are surely weaker than if I *actually had* contracted to do *A* upon your doing *B*.

Further, in the case of the hypothetical consent to a Kaldor-Hicks move that flows from probabilistic information, it is not clear why hypothetical consent should have *any* moral force. Suppose a person (*P*) is, ex post, a loser from the Kaldor-Hicks move, but as a result of risk or uncertainty, *P*'s ex ante assessment (if *P* were to make such an assessment) would be that the move would produce an expected gain for him. *P* does not actually consent to the move, but he would do so if the question were raised in a scenario of social contracting. Why should this sway government in favor of the move? Note that *P*'s hypothetical consent to the move is partly uninformed; he would approve it only because he is ignorant of the fact that it will end up harming him. In short, the best version of the consent-based argument for Kaldor-Hicks appeals not to actual consent, nor to fully informed hypothetical consent, but to partly uninformed hypothetical consent, and thus is hardly compelling, even in the empirically contingent case where the partial information available to each actor makes the project seem an expected gain for her.

Sixth, the Kaldor-Hicks standard is sometimes defended as a workable decision procedure. The idea is that Kaldor-Hicks offers a clearer and easier way for agencies, courts, applied economists, and/or legal scholars to evaluate governmental choices (as compared to direct evaluation of governmental choices in light of a full social welfare function) and that Kaldor-Hicks efficiency tracks the true social welfare function sufficiently well to be a viable proxy.⁵⁰ This is

49. See, e.g., Heidi M. Hurd, *Justifiably Punishing the Justified*, 90 MICH. L. REV. 2203, 2305 (1992) (contrasting moral force of actual and hypothetical consent).

50. See Arnold C. Harberger, *Three Basic Postulates for Applied Welfare Economics: An Interpretive Essay*, 9 J. ECON. LIT. 785, 796 (1971) (presenting three "basic postulates")

quite plausible.⁵¹ But it fails to answer the question under consideration now: is Kaldor-Hicks a *normative* (specifically, a moral) criterion?⁵²

A theory of regulation starts from a normative framework with reference to which regulatory policies, choices, institutions, and decision procedures can be assessed. In order to know whether Kaldor-Hicks is a “good” or “viable” or “appropriate” decision procedure—that is, in order to assess the claim that it generally produces “good” or “good enough” outcomes, that its decisional “costs” are “low,” that it is unlikely to be “misused” by practitioners, and so on—we need criteria for appraising the goodness and badness of outcomes, the existence and level of costs, and the fact and extent of misuse. And even if this is denied—even if we somehow can intuitively “know” that Kaldor-Hicks is a normatively appropriate decision procedure without knowing what the criteria for normative appropriateness are—the point remains that justified decision procedures and normative criteria are distinct. A given standard can constitute a justified decision procedure without constituting a normative criterion, and vice versa. Neoclassicism is the view that posits Kaldor-Hicks efficiency as one (if not the only) basic normative criterion. The question at hand is whether Kaldor-Hicks efficiency indeed has this status; and to characterize it as a viable decision procedure is simply to switch debates.

Seventh and finally, the defense of the Kaldor-Hicks standard is sometimes married to a certain moral skepticism, in particular a skepticism about the possibility of interpersonal welfare comparisons.⁵³ This was certainly true of the original defense of the standard, set forth by the economists Kaldor, Hicks, and Scitovsky.⁵⁴ They proposed the hypothetical-compensation standard in a milieu where overall well-being was thought to be indeterminate or unknowable,⁵⁵ and as a way to rank Pareto-noncomparable states without making interpersonal welfare comparisons. But as I shall argue at greater

as the answer to a need for a set of standards, and arguing they provide a basis for “a professional consensus on procedures”).

51. See, e.g., Adler & Posner, *supra* note 28, at 194-243 (defending cost-benefit analysis as a justified decision procedure in light of the moral criterion of overall well-being).

52. On the difference between decision procedures and moral criteria, see DAVID O. BRINK, MORAL REALISM AND THE FOUNDATIONS OF ETHICS 216-17 (1989).

53. It is unclear to me whether Richard Posner is still committed to the Kaldor-Hicks standard, but if so, that commitment is now wedded to a kind of moral skepticism. See RICHARD A. POSNER, THE PROBLEMATICS OF MORAL AND LEGAL THEORY *passim* (1999) (attack on moral theorizing); *id.* at 235 (equating “normative economics” and “cost benefit analysis”).

54. See Hicks, *supra* note 8; Kaldor, *supra* note 8; Tibor de Scitovsky, *A Note on Welfare Propositions in Economics*, 9 REV. ECON. STUD. 77 (1941).

55. See Lionel Robbins, *Interpersonal Comparisons of Utility: A Comment*, 48 ECON. J. 635 (1938) (articulating a skeptical view about interpersonal welfare comparisons).

length below, the criterion of overall well-being is neither indeterminate nor unknowable. Modern theoretical welfare economists, unlike Kaldor and Hicks and their contemporaries, are generally not skeptics about the possibility of interpersonal comparisons.⁵⁶

In any event, even if such skepticism were warranted, the leap to the Kaldor-Hicks criterion would be unjustified. We would still need an affirmative argument why, given two Pareto-noncomparable states, the one that is Kaldor-Hicks efficient is better. I have tried to show that no such argument is forthcoming; thus skepticism about interpersonal comparisons, if warranted, would simply warrant us in a general skepticism about the normative comparability of Pareto-noncomparable states via judgments of “overall well-being” or “Kaldor-Hicks efficiency”—or anything else.

My argument against the moral significance of Kaldor-Hicks efficiency concludes here. That Kaldor-Hicks lacks moral importance is, as I have said, hardly a novel point. It is a point ably articulated in the legal literature.⁵⁷ But scholarly appeals to the Kaldor-Hicks standard—explicit appeals or, as in the case of some neoclassicists, implicit appeals—have not ceased, and so the point bears the emphasis and elaboration that I have given it. Further, it should be underlined that someone who rejects the moral importance of Kaldor-Hicks efficiency (as I do) need not reject it for all the various reasons presented by its past critics. The argument against Kaldor-Hicks presented here is, in significant ways, different from and more modest than the arguments previously advanced by scholars who oppose that standard or, more generally, law and economics.

2. *Some Spurious Criticisms of Kaldor-Hicks Efficiency*

Note what I have *not* said in criticism of Kaldor-Hicks. To begin with, I have not criticized Kaldor-Hicks for sometimes leading to an indeterminate ranking of social states. This is the so-called Scitovsky paradox: it is possible for one social state (S_1) to be Kaldor-Hicks efficient relative to another (S_2), while at the same time S_2 is Kaldor-Hicks efficient relative to S_1 .⁵⁸ The Scitovsky paradox is closely re-

56. See, e.g., BOADWAY & BRUCE, *supra* note 29, at 137-69. The so-called Bergson-Samuelson social welfare function is routinely employed by modern welfare economists as a criterion for evaluating social states; it assumes interpersonally comparable utilities. See *id.* But see Robert A. Pollak, *Welfare Comparisons and Situation Comparisons*, 50 *J. ECONOMETRICS* 31, 31 (1991) (claiming that “most economists think interpersonal comparisons are nonsense”).

57. See sources cited *supra* notes 30-33.

58. See Scitovsky, *supra* note 54. For textbook discussions, see, for example, BOADWAY & BRUCE, *supra* note 29, at 96-101; YEW-KWANG NG, *WELFARE ECONOMICS: INTRODUCTION AND DEVELOPMENT OF BASIC CONCEPTS* 59-66 (1980). A more precise formulation of the Scitovsky paradox is as follows. Given two states, S_i and S_j , where S_i is the status quo and S_j is an alternative, let us say that the Winners in S_i are those persons who are better off in S_i than S_j ; that the Losers in S_i are those persons who are worse off in S_i than S_j ; that

lated to another phenomenon, often discussed in the legal literature—the fact that P 's “willingness to pay” to gain a good does not necessarily equal what he would be “willing to accept” in compensation for its loss.⁵⁹ That is, a person's preferences as between various goods can change as her total wealth changes. And sometimes the critics of Kaldor-Hicks have argued that, because of the Scitovsky paradox, Kaldor-Hicks cannot be a true moral criterion.⁶⁰ But this criticism is unpersuasive.

The Scitovsky paradox means that the Kaldor-Hicks criterion does not enable a “complete” ranking of social outcomes such that each outcome is better, worse, or precisely equal to every other. But so what? Genuine moral criteria do not necessarily enable complete rankings of social outcomes. For example, the criterion of overall well-being does not: if one social state (S_1) is a little more pleasurable for one person (P_1) than a second state (S_2), but S_2 allows a second person (P_2) to accomplish his professional goals a little better than S_1 , then S_1 may be neither better, nor worse, nor precisely equal to S_2 in light of overall well-being. It may instead be welfare-incomparable. Similarly, if S_1 decreases the cumulative deviation from average welfare relative to S_2 , but S_2 decreases the cumulative deviation from the welfare of the richest person relative to S_1 , then S_1 and S_2 may be neither better, nor worse, nor precisely the same with respect to the criterion of equality.⁶¹ Indeed, the point that true normative criteria need not enable complete rankings of states should have been obvious from an examination of the Pareto crite-

the Winners in S_j are those persons who are better off in S_j than S_i ; and that the Losers in S_j are those persons who are worse off in S_j than S_i . Note that the Winners in S_i are just the same persons who are the Losers in S_j , and that the Losers in S_i are just the same persons who are the Winners in S_j . So there are really only two relevant categories of persons, who can be referred to as the Winners, those who benefit from a change in the status quo (the Winners in S_j /Losers in S_i) and the Losers, those who lose from a change in the status quo (Losers in S_j /Winners in S_i).

What Scitovsky pointed out was this: It is possible for states S_1 and S_2 to exist, with S_1 the status quo and S_2 the alternative, such that (1) there is a reallocation of goods in S_2 from the Winners to the Losers that makes at least one person better off in S_2 and no one worse off, as compared to S_1 ; and yet (2) there is a reallocation of goods in S_1 from the Losers to the Winners that makes at least one person better off in S_1 and no one worse off, as compared to S_2 . See *id.*

59. See generally Elizabeth Hoffman & Matthew L. Spitzer, *Willingness to Pay vs. Willingness to Accept: Legal and Economic Implications*, 71 WASH. U. L.Q. 59 (1993).

60. For an example of this criticism within the legal literature, see Coleman, *supra* note 31, at 518-20; within the theoretical literature on welfare economics, see BOADWAY & BRUCE, *supra* note 29, at 96-101; within the literature on applied welfare economics and cost-benefit analysis, see PEARCE & NASH, *supra* note 36, at 28-31; and within the philosophical literature, see HAUSMAN & MCPHERSON, *supra* note 9, at 96.

61. See *infra* text accompanying notes 158-61 (arguing that outcomes can be incomparable with respect to welfare and other moral criteria, and citing sources); see also LARRY S. TEMKIN, *INEQUALITY* 19-52 (1993) (suggesting that cumulative deviation from average welfare and cumulative deviation from the welfare of the richest person are, *inter alia*, possible conceptions of equality).

rion; where P_1 is better off in S_1 , and P_2 is better off in S_2 , neither state is better than, worse than, or precisely equal to the other with respect to Pareto-efficiency, since some S_3 can be Pareto-efficient relative to S_1 but not S_2 .

Nor have I tasked Kaldor-Hicks efficiency for running afoul of the “incommensurability” of diverse goods. Kaldor-Hicks efficiency and its cousin, cost-benefit analysis, have been roundly criticized for reducing all the various impacts that governmental choices can have to dollar- or resource-equivalents. These criteria involve the “pricing” of physical safety, friendship, environmental preservation, sexual fulfillment, and other such “priceless” goods, or so the criticism goes.⁶² For reasons I have articulated at length elsewhere,⁶³ I think the “incommensurability” criticism of Kaldor-Hicks efficiency or of cost-benefit analysis fails. If Kaldor-Hicks efficiency did track something of basic importance about outcomes—long-run Paretianism, expected Paretianism, consent, or whatever—neither incommensurability nor the Scitovsky paradox would preclude counting Kaldor-Hicks efficiency as a normative criterion and operationalizing it through the cost-benefit procedure.

Finally, I have not faulted Kaldor-Hicks efficiency for running afoul of “rights” (deontological criteria), equality, perfectionist values, overall well-being, or any other normative criteria distinct from Kaldor-Hicks that the critic might posit. To be sure, one state (S_1) can be Kaldor-Hicks efficient relative to another (S_2) even though S_2 is better in light of rights, equality, perfectionist values, and so on. Then again, S_3 can be better than S_4 in light of rights, even though S_4 is better in light of equality. And S_5 can be better than S_6 in light of equality, even though S_6 is better in light of overall well-being. Distinct moral criteria can *conflict*; that is exactly what makes them distinct. Unless our moral framework is assumed to be a monistic one—and why should we assume that?⁶⁴—the fact that a proposed moral

62. See, e.g., ELIZABETH ANDERSON, VALUE IN ETHICS AND ECONOMICS 190-216 (1993); MARTHA C. NUSSBAUM, *Plato on Commensurability and Desire*, in LOVE'S KNOWLEDGE: ESSAYS ON PHILOSOPHY AND LITERATURE 106 (1990); MARGARET JANE RADIN, CONTESTED COMMODITIES 115-22 (1996); Lester B. Lave, *Benefit-Cost Analysis: Do the Benefits Exceed the Costs?*, in RISKS, COSTS, AND LIVES SAVED: GETTING BETTER RESULTS FROM REGULATION 104, 113 (Robert W. Hahn ed., 1996); Richard H. Pildes & Cass R. Sunstein, *Reinventing the Regulatory State*, 62 U. CHI. L. REV. 1, 64-66 (1995); Cass R. Sunstein, *Incommensurability and Valuation in Law*, 92 MICH. L. REV. 779, 782-85 (1994).

63. See Matthew Adler, *Incommensurability and Cost-Benefit Analysis*, 146 U. PA. L. REV. 1371, 1401-17 (1998).

64. See, e.g., MICHAEL STOCKER, PLURAL AND CONFLICTING VALUES (1990) (arguing against monism).

65. “Outcomes” here are total world-histories. If outcomes are taken instead to mean discrete states of affairs, then the preference-based view claims that P is *intrinsically* benefited by an outcome (thus defined) if, and only if, he prefers it. On this ambiguity in the notion of an outcome or state, see Adler & Posner, *supra* note 28, at 202 n.97.

criterion can conflict with another, accepted criterion, is no grounds for rejecting the proposal.

The fact that I do not fault Kaldor-Hicks efficiency on the grounds of incompleteness (Scitovsky reversals), incommensurability, or potential conflict with other criteria will become significant when I present my own theory of regulation, the welfarist theory. Such criticisms could be leveled against welfarism, just as they have been leveled against Kaldor-Hicks; but they are effective against neither.

C. *The Preference-based View of Well-being*

So much for Kaldor-Hicks efficiency, the first component of neoclassicism. What about the second component, the preference-based view of well-being? There are strong grounds for rejecting a straight preference-based view. It is less clear whether the best alternative is (1) a modified preference-based view, or (2) an objective-list view with some minimal role for preferences. I tend to think that the most persuasive theory of well-being is the first, a modified preference-based view. This is less of a departure from neoclassicism than the second alternative, objectivism, but the difference will be sufficiently significant for regulatory purposes that it merits delineation.

1. *Problems with the Preference-based View*

What's wrong with the straight preference-based view of well-being? The view is that a person (P) is benefited by one outcome (S_1) as compared to another (S_2), just in case P prefers S_1 to S_2 .⁶⁵ One large problem is that preferences can be motivated by a range of considerations, including but not limited to P 's welfare.⁶⁶ A preference is simply a *ranking* of world-states or, perhaps, a ranking that is conceptually connected to the actions of the person who holds the preference.⁶⁷ In either case, P 's grounds for the ranking can be anything that motivates him to count one outcome as better than the other.

66. This difficulty with the straight preference-based view of well-being is widely recognized within the philosophical literature on well-being. See SUMNER, *supra* note 9, at 134-35; John Broome, *Choice and Value in Economics*, 30 OXFORD ECON. PAPERS 313 (1978); Allan Gibbard, *Interpersonal Comparisons: Preference, Good, and the Intrinsic Reward of a Life*, in FOUNDATIONS OF SOCIAL CHOICE THEORY 165, 173-75 (Jon Elster & Aanund Hylland eds., 1986); Mark Carl Overvold, *Self-Interest and the Concept of Self-Sacrifice*, 10 CANADIAN J. PHIL. 105 (1980); Amartya K. Sen, *Rational Fools: A Critique of the Behavioral Foundations of Economic Theory*, 6 PHIL. & PUB. AFF. 317 (1977); David Sobel, *On the Subjectivity of Welfare*, 107 ETHICS 501 (1997).

67. On the nature of a "preference" and the possible conceptual connection between preference and choice, see, for example, S.L. HURLEY, *NATURAL REASONS: PERSONALITY AND POLITY* 55-83 (1989); PREFERENCES (Christoph Fehige & Ulla Wessels eds., 1998); Richard Arneson, *Liberalism, Distributive Subjectivism, and Equal Opportunity for Welfare*, 19 PHIL. & PUB. AFF. 158, 161-64 (1990); Arthur Ripstein, *Preference, in VALUE, WELFARE, AND MORALITY* 93-111 (R.G. Frey & Christopher W. Morris eds., 1993).

For example, P might rank S_1 over S_2 on moral grounds. Thus he may prefer S_1 because he believes that equality is of prime moral significance and S_1 advances the value of equality better than S_2 , or because he believes that the environment is intrinsically valuable and S_1 is better than S_2 in that regard, or because S_2 involves some action of his that he takes deontological constraints to prohibit. Alternatively, P might rank S_1 over S_2 by virtue of some relationship he has. Thus, he may believe that S_1 is better for the welfare of some friend or relation of his and, although P is personally indifferent between the two outcomes, choose S_1 out of loyalty. Or, the ranking might be grounded in a role of P s: perhaps P is an American, or a Jew, or a teacher, and it is constitutive of that role to prefer S_1 to S_2 . Finally, P might rank S_1 over S_2 for reasons of nonmoral goodness: it does not matter to his life whether S_1 or S_2 occurs, but S_1 is better, say, in light of aesthetic values.⁶⁸

There are surely *some* kinds of motivating reasons such that, if P is motivated to rank S_1 over S_2 by virtue of these kinds of reasons, he is indeed typically made better off by S_1 . Call such reasons welfare-productive. We might agree that self-interested reasons are welfare-productive: if P prefers S_1 to S_2 on the grounds of his own welfare, then he is typically better off if S_1 occurs. We might also think that reasons concerning certain objective welfarist goods are welfare-productive: if P prefers S_1 to S_2 on the grounds that S_1 involves more fun for him, or a more genuine accomplishment, or a deeper involvement with important others, then P is typically better off if S_1 occurs.

However one delineates the category of welfare-productive reasons, it is implausible that *all* reasons motivating preferences are welfare-productive. The straight preference-based view insists that moral reasons, relationship-based reasons, role-relative reasons, reasons of nonmoral goodness, and all other considerations that might motivate P to prefer or rank S_1 over S_2 are welfare-productive. But why think that? To begin with, it is counterintuitive. More seriously, it is ungrounded; there is nothing in the possible connection between preference and choice, or in the general desiderata of a theory of well-being, to move us even *prima facie* towards thinking that all preference-motivating reasons are welfare-productive. As for the connection between preference and choice, people can certainly choose options that have no personal benefit or are personally detrimental. As for the general desiderata of a theory of well-being, it is certainly plausible that preferences are a *necessary* condition for a welfare improvement (P cannot be made better off by S_1 over S_2 unless, *inter alia*, he prefers S_1 to S_2), but it is a logical mistake to

68. The possible disjunction between welfare and preference satisfaction in the kind of cases just described is discussed at length in David Sobel, *Well-Being as the Object of Moral Consideration*, 14 *ECON. & PHIL.* 249 (1998).

leap from this plausible point to the conclusion that preferences are both necessary and sufficient for a welfare improvement (P is made better off by S_1 over S_2 just in case he prefers S_1 to S_2), which is what the straight preference-based view of welfare says.⁶⁹

A second problem with the straight preference-based view of well-being is that preferences can be uninformed, unreflective, nonautonomous, or otherwise nonideal.⁷⁰ A person may, for example, prefer smoking to abstinence; were she fully informed about the damage that smoking does to her internal organs, she would prefer abstinence. Another person may prefer building a house on the mountain-side lot that she owns to leaving it bare. She knows that the mountainside is part of a rare, old-growth forest, but not having fully deliberated on that point, she still prefers to build. Only if the information about the forest were integrated with her other, background information and with her background commitments, values, and goals, would she come to prefer an undeveloped lot.

One standard view in the philosophy of well-being is that "ideal" conditions for preference formation are simply the conditions of full information and deliberation,⁷¹ but it is not clear whether this is right.⁷² For example, suppose a person (P), who was horribly abused as a child, prefers to torture cats. He retains this preference notwithstanding full information about the pain that his actions cause the suffering animal, and notwithstanding all the deliberation of which he is capable. Yet it seems counterintuitive that the world is made better, even *pro tanto*, when P tortures a cat. To resist this conclusion, one might characterize P 's preferences (albeit fully informed and deliberate) as nonautonomous, in the sense that he lacks the capacity to revise his preferences, or in the sense that the set of preferences which are possible for him is not sufficiently large, or in the sense that his preferences are rooted in a particular sort of upbringing, or in some such related sense.

2. *Alternative Theories of Well-Being*

I see two theories (more precisely, families of theories) of welfare that are responsive to the problems with the straight preference-based view just described and yet retain enough of that view to re-

69. See Adler & Posner, *supra* note 28, at 197-204 (distinguishing between the claim that preference satisfaction is necessary for welfare and the claim that preference satisfaction is both necessary and sufficient for welfare, and defending the former).

70. See, e.g., SUMNER, *supra* note 9, at 156-71.

71. See David Sobel, *Full Information Accounts of Well-Being*, 104 ETHICS 784, 792 n.15 (1994) (noting the prevalence of full-information accounts of well-being within philosophical scholarship and citing sources).

72. See AMARTYA SEN, ON ETHICS AND ECONOMICS 45-46 (1987); SUMNER, *supra* note 9, at 162-71.

main plausible accounts of what makes a life go well.⁷³ The hedonic view of welfare—equating welfare with pleasurable sensations and the absence of painful sensations—“solves” these problems at the significant cost of excluding preferences from the picture entirely. Even if hedonism were amended to make joint reference to pleasure/pain *and* preference satisfaction, the problem would remain that we can prefer for ourselves, and benefit from, much more than the presence of pleasure and the absence of pain. Objectivism that counts hedonic sensations as one entry on a larger list of objective goods⁷⁴ is surely an improvement over straight hedonism, but it falls to the criticism that no amount of objective goodness in one state (S_1) over another (S_2) can make S_1 better for a person (P) unless P prefers (or at least comes to prefer) S_1 .

This leaves the following candidates.

Sophisticated Preferentialism: P is better off with S_1 , as compared to S_2 , just in case (1) P prefers S_1 over S_2 ; and (2) P would prefer S_1 over S_2 under ideal conditions; and (3) P 's preference and ideal preference are suitably restricted.

Sophisticated Objectivism: P is better off with S_1 , as compared to S_2 , just in case (1) P does better in S_1 , in light of the totality of objective welfare goods (friendship, accomplishment, community involvement, fun, pleasure, and so on); and (2) P prefers S_1 over S_2 .

How exactly the notion of “ideal conditions” should be specified is a matter for further debate among sophisticated preferentialists. A parallel point can be made about the goods identified as “objective welfare goods” by sophisticated objectivists, and about the grounds for thus identifying them. This is the sense in which these are families of welfare theories rather than specific theories. Note that both kinds of theories are sensitive to the criticism I leveled above against simple objectivism—namely, P cannot be made better off in the teeth of his actual preferences. This is the key insight behind the straight preference-based view. It is incorporated into sophisticated preferentialism and sophisticated objectivism by making P 's actual preference for an outcome a necessary (but not sufficient) condition for that outcome to improve his well-being.

What about the choice *between* sophisticated preferentialism and sophisticated objectivism? Both are plausible, but which is the uniquely *right* theory or kind of theory? This is a difficult question. Sophisticated preferentialism, as here articulated, contains a large

73. See generally JAMES GRIFFIN, WELL-BEING 7-72 (1986) (discussing different theories of well-being, specifically hedonic, preference-based, and objectivist views); DEREK PARFIT, REASONS AND PERSONS 493-502 (1984) (same); SUMNER, *supra* note 9 (same).

74. For plausible examples of such lists, see JOHN FINNIS, NATURAL LAW AND NATURAL RIGHTS 85-90 (1980); GRIFFIN, *supra* note 73, at 67-68; and GEORGE SHER, BEYOND NEUTRALITY: PERFECTIONISM AND POLITICS 199-201 (1997).

and embarrassing lacuna: the failure to state just how preferences should be “restricted” such that thus-restricted preferences, and only those, are welfare productive. Filling the gap by requiring that preferences be “self-interested” or that they “concern *P*’s life” would be circular; these formulations, in effect, define the class of “restricted” preferences by reference to their effect on welfare, yet it was the very concept of welfare that sophisticated preferentialism was trying to explain through the notion of restricted preferences.

Mark Overvold has developed a noncircular account of restricted preferences.⁷⁵ Roughly, *P*’s preference for an outcome is welfare-productive for *P* if the outcome entails *P*’s existence. But this runs afoul of the objection that various morally required outcomes can entail *P*’s existence, and be preferred by him on moral grounds, without benefiting him.⁷⁶ (Imagine the case where *P*, who has committed wrong, prefers that he suffer because he believes retributivism about punishment to be true.) Other philosophers have, in less detail, proposed accounts of preference restriction different from Overvold’s, but these accounts, too, have troubling and maybe fatal flaws.⁷⁷

On the other hand, sophisticated objectivism may not be able to avoid the problem of restricting preferences. That problem is superficially avoided through the incorporation of objective goods, where each good constitutes a criterion for ranking outcomes and, for each pair of outcomes, specifying those persons who have fared better or worse as between the two. (So, for example, the value of “fun” would tell us whether a given outcome is more fun for anyone and, if so, for whom.) But the avoidance may be superficial. If the objectivist simply gives us a list of objective goods, we can protest that a mere list is ad hoc. What are the grounds for placing various goods on the list?⁷⁸ And if the objectivist tries to ground his list in collective or normal preferences of some sort (as many working in value-theory now try to do), we are back to the problem of restriction.

Quite apart from all this, sophisticated objectivism is vulnerable to the criticism that it is conformist. If “objective values” for a given *P*

75. See Mark Carl Overvold, *Morality, Self-Interest, and Reasons for Being Moral*, 44 PHIL. & PHENOMENOLOGICAL RES. 493, 499-501 (1984); Mark Carl Overvold, *Self-Interest and Getting What You Want*, in THE LIMITS OF UTILITARIANISM 186 (Harlan Miller & William H. Williams eds., 1982); Overvold, *supra* note 66, at 117-18 n.10.

76. See Sobel, *supra* note 66, at 505-06 (criticizing Overvold’s theory on this and other grounds).

77. See *id.* (generally discussing and criticizing attempts to “restrict” preference, i.e., to identify some subset of welfare-productive preferences).

78. L.W. Sumner has made this objection:

The particular items on Finnis’s list [of objective goods] do not matter for our present purposes; what does matter is that he has no account to offer of what makes something (anything) a source of our welfare—what gains it a place on the list—if this does not depend on our attitudes or concerns.

SUMNER, *supra* note 9, at 46.

are specified as the things that *P* would prefer under ideal conditions, then sophisticated objectivism is not an alternative to preferentialism. But if “objective values” are specified more narrowly, then (seemingly) it becomes impossible for *P* to benefit by an idiosyncratic or contrarian way of life, since such a life is one that is good for *P* only in being endorsed by him and in no other way.

I tend to think that sophisticated preferentialism is on balance a better welfare view than sophisticated objectivism, and that the problem of preference restriction—which clearly bedevils the first view, if not the second—will eventually be solved. Further, preferentialist views are much more widely adopted by modern philosophers and other contemporary scholars doing rigorous work on the nature of well-being.⁷⁹ Finally, sophisticated preferentialism is closer to neoclassicism than sophisticated objectivism, and it thus represents a less dramatic change in prevailing theories of regulation. For all these reasons, the remainder of this Article is written on the premise that sophisticated preferentialism (not sophisticated objectivism) is the correct account of well-being. That premise will be of some significance in Part III, where I criticize so-called proceduralist theories of regulation; and it will be of particular importance in Part IV, where I turn from my criticism of neoclassicism and proceduralism to defend my own, novel theory of regulation, which I call “welfarism.”

III. PROCEDURALIST THEORIES OF REGULATION

A. *Characterizing Proceduralism*

I use the term “proceduralist” to mean a moral theory of regulation which accords intrinsic significance, not merely instrumental significance, to the fact that a regulatory agency followed or failed to follow some specified type of procedure. In other words, a proceduralist theory stipulates that the moral status of a regulatory choice is a function of criteria $X_1, X_2 \dots X_n$, where at least one X_i is a procedural criterion in that it refers not to the legal content of the directives issued by regulators, nor to the broader social context in which those directives are issued, but specifically to the decisionmaking *processes* that the regulators traversed (or failed to traverse) in issuing the directives. The most significant body of theoretical work on regulation, other than neoclassicism, consists of theories that are (or at least are plausibly understood as) proceduralist. Under this rubric I include traditional and modern variants of pluralism, civic republican theo-

79. See Sobel, *supra* note 71, at 792 n.15 (“[A] truly impressive and diverse list of contemporary ethicists have found [some kind of desire-based or preference-based] account of well-being congenial.”).

ries, and negotiated rulemaking or collaborative governance theories.⁸⁰

The pluralists, the civic republicans, and the collaborative-governance theorists clearly direct their attention to the procedures by which regulatory policies are enacted and advocate certain procedural alternatives while criticizing others. This common feature of these theories does not yet demonstrate them to be proceduralist—as I have defined that term—because it is possible that their concern for procedural forms is instrumental, not intrinsic. The pluralist might simply think that a process of self-interested bargaining among groups is the best way to effect the substantive moral criteria that govern regulatory choices: criteria such as welfare maximization, the equalization of welfare or the opportunity for welfare, and the promotion of the environment or other perfectionist values. Similarly, the civic republican's commitment to public-spirited deliberation about the common good might be grounded in the contingent and empirical claim that deliberation produces “good” or “right” laws and other regulatory directives, with “good” or “right” specifiable solely in terms of the legal content of the directives plus general social context, and not in terms of the decisionmaking processes that were used to formulate the directives.⁸¹ And the collaborative-governance theorist might advocate regulatory procedures characterized by “[a] problem-solving orientation,” “[p]articipation by interested and affected parties in all stages of the decision making process,” “[p]rovisional solutions,” “[a]ccountability that transcends traditional public and private roles in governance,” and “[a] flexible, engaged agency”⁸²—to use Jody Freeman's list of procedural desiderata—on purely instrumental grounds.

But there is significant textual evidence that at least some pluralists, civic republicans, and collaborative-governance theorists do have a noncontingent and nonempirical attachment to regulatory procedures—in short, that they are proceduralist, as I have defined that term.⁸³ Note further that proceduralism constitutes a distinctive

80. See sources cited *supra* notes 3-7.

81. One author has noted:

[It is open to] question whether civic republicanism defines the public good procedurally or substantively. If defined substantively, the common good is akin to a right answer, and it implies a politics that emphasizes deliberation in order to ensure higher-quality decisionmaking. If defined procedurally, the public good . . . implies a politics that emphasizes participation so that each citizen has the opportunity to experience self-government.

Note, *supra* note 6, at 1404.

82. Freeman, *supra* note 7, at 22-23.

83. See, e.g., Freeman, *supra* note 7, at 27 (“[P]articipation in regulatory problem solving by interested and affected parties has an independent, democratic value.”); Gey, *supra* note 6, at 810 (claiming that civic republicans define virtuous outcomes as those that result from selfless, collective deliberation); Seidenfeld, *supra* note 4, at 1514 (“[Under civic re-

type of moral theory of regulation, while instrumentalism does not. Instrumentalism about a given procedure (to be plausible) presupposes some prior and independent theory of the criteria $X_1, X_2 \dots X_n$ applicable to regulatory choices; for it would only be plausible to say that a given procedure has the empirical tendency to produce outcomes that are better with respect to a *given* set of criteria. It would not be plausible to suggest that a given procedure has the tendency to produce outcomes that are better with respect to the correct set of criteria *whatever* that set happens to be.

For these reasons, I will generally focus in this section upon the proceduralist construal of pluralism, civic republicanism, and collaborative-governance theory. Bargaining, deliberation, or collaboration *can* be defended instrumentally, but that possibility will be kept in abeyance until it becomes clear why proceduralism fails. Further, I will launch a *global* attack on proceduralism, an attack which is meant to target all its variants, including the pluralist variant that takes bargaining to be of intrinsic importance, the civic-republican variant that takes deliberation to be of intrinsic importance, and the collaborative-governance variant that takes a certain kind of collaborative engagement by the regulatory agency and interested parties to be of intrinsic importance. Readers of the legal scholarship on regulation are well accustomed to civic-republican arguments against pluralism,⁸⁴ to collaborative-governance arguments against civic-republicanism,⁸⁵ and to other scholarship that criticizes one or more procedural forms only to conclude by favoring another. My approach will be less discriminating and, therefore, less familiar. Proceduralists of all stripes have made a common mistake, and their theories are equally flawed, since the moral criteria governing regulatory choice do *not* include procedural criteria—or so I will now try to show.

B. *The Defense Based on Moral Skepticism*

What's wrong with proceduralism? Let me flip the question around: How might proceduralism, of any kind, be defended? First, the proceduralist might be a *moral skeptic*. She might think that

publicanism] government's primary responsibility is to enable the citizenry to deliberate about altering preferences and to reach consensus on the common good."); *id.* at 1535 ("Some commentators have even suggested that participation in determining the identity of the political community is itself a good, thereby rendering civic republicanism inherently as well as instrumentally valuable.") (citing sources); SUNSTEIN, *supra* note 5, at 35 ("A different conception of autonomy places an emphasis on the freedom of collectivities or communities—a freedom embodied in decisions, reached by the citizenry as a whole, about what courses to pursue. . . . An important form of freedom may consist in precisely these processes of collective self-determination.").

84. See, e.g., Seidenfeld, *supra* note 4, at 1533-34.

85. See, e.g., Freeman, *supra* note 7, at 19-21.

regulatory choices cannot be evaluated as morally better or worse, right or wrong—that where the regulator has the option of choosing one of two outcomes (O_1 or O_2), any statement to the effect that the regulator is truly morally obliged to choose O_1 (because O_1 maximizes welfare, or equalizes welfare, for example) is nonsense. There is certainly a flavor of skepticism in some versions of pluralism,⁸⁶ which originated at a time when logical positivism⁸⁷ and legal realism⁸⁸ were ascendant, and when many philosophers and other intellectuals indeed denied the existence of moral truth. The argument from skepticism to pluralism that pluralists (at least traditional ones) sometimes seem to advance goes roughly like this:

- (1) there is no objective moral standard by which to assess governmental choices;
- (2) rather, statements purportedly describing choices as better or worse, right or wrong, simply express the subjective preferences of the speaker;
- (3) more generally, human actions can be fully explained by reference to the preferences and beliefs of actors, with no need to invoke moral criteria or anything else beyond preferences as the driving force behind actions; and therefore
- (4) all that is true, and all that needs to be true, of governmental choices is that they result from a process of bargaining between self-interested actors, each seeking to maximize her preferences given her beliefs.⁸⁹

This argument is both unsound and invalid. It is unsound because the first two premises are untrue. There are objective moral standards applicable to governmental choices and other human actions; statements attributing moral obligations, permissions, and so forth, to governmental actors are meaningful, sometimes true, and need *not* merely express the preferences of the speaker. A persuasive case against moral skepticism, and in favor of moral cognitivism, has been

86. As Richard Stewart observes:

Implicit in [the interest-group model of administrative law] is the assumption that there is no ascertainable, transcendent 'public interest,' but only the distinct interests of various individuals and groups in society. Under this assumption, legislation represents no more than compromises struck between competing interest groups. . . . In the extreme form of this view, there is no objective, independent yardstick by which one can measure the content of compromise; compromises are legitimated by the process of their negotiation. . . . Pluralist political theory may be regarded as a translation into collective terms of the principle of subjectivity of individual values.

Stewart, *supra* note 3, at 1712 & n.206 (citing ARTHUR FISHER BENTLEY, *THE PROCESS OF GOVERNMENT* (1908), and DAVID B. TRUMAN, *THE GOVERNMENTAL PROCESS* (1951)).

87. See, e.g., ALFRED J. AYER, *LANGUAGE, TRUTH AND LOGIC* (1936).

88. See generally LAURA KALMAN, *LEGAL REALISM AT YALE, 1927-1960* (1986).

89. See Stewart, *supra* note 3, at 1671-76; see also ARTHUR FISHER BENTLEY, *THE PROCESS OF GOVERNMENT* (1908); DAVID B. TRUMAN, *THE GOVERNMENTAL PROCESS* (1951).

presented by a number of contemporary philosophers working in metaethics, including Michael Moore,⁹⁰ David Brink,⁹¹ Michael Smith,⁹² Geoffrey Sayre-McCord,⁹³ John McDowell,⁹⁴ and others. Cognitivism is probably now the dominant (if not exclusive) position in philosophy;⁹⁵ the skeptical consensus characteristic of the era of logical positivism is surely long past. But even if the demise of moral skepticism were unwarranted, and the initial premises of the argument just sketched were true, the problem would remain that the inference from those two premises plus the third to the pluralist-proceduralist conclusion is mistaken. The three premises merely imply the following: (4)* all that is true of governmental choices is that they result from a process of bargaining by self-interested actors. The premises do not additionally imply that governmental choices *need to*, *ought to*, or *should* result from a bargaining process. In fact, the premises imply the opposite: if in general objective moral criteria do not exist, then in particular the objective moral criteria that might require a bargaining process do not exist, and the purportedly objective statement that governmental choices *need to*, *ought to*, or *should* result from such a process is nonsense. The only kind of statement, objectively, that the moral skeptic-cum-pluralist is permitted by her own skepticism to utter is a positive statement about the kinds of governmental processes that do or will occur, namely (4)*.⁹⁶ Normative pluralism grounded upon skepticism is self-defeating.

More generally, any kind of proceduralism founded upon skepticism is self-defeating. Skepticism denies that moral truths exist, and it construes moral claims as playing the "expressive" role of articulating the speaker's preferences, or some other such non-truth-stating role.⁹⁷ But then the claim that governmental actors are morally

90. See, e.g., Michael Moore, *A Natural Law Theory of Interpretation*, 58 S. CAL. L. REV. 277 (1985).

91. See BRINK, *supra* note 52.

92. See MICHAEL SMITH, *THE MORAL PROBLEM* (1994).

93. See Geoffrey Sayre-McCord, *Moral Theory and Explanatory Impotence*, in *ESSAYS ON MORAL REALISM* 256 (Geoffrey Sayre-McCord ed., 1988).

94. See John McDowell, *Values and Secondary Qualities*, in *MORALITY AND OBJECTIVITY* 110 (Ted Honderich ed., 1985).

95. See generally Stephen Darwall et al., *Toward Fin de siècle Ethics: Some Trends*, in *MORAL DISCOURSE AND PRACTICE: SOME PHILOSOPHICAL APPROACHES* 3 (Steven Darwall et al. eds., 1997) (providing overview of contemporary metaethics, in particular the debate between cognitivists and noncognitivists). See also BRINK, *supra* note 52 (same); DAVID MCNAUGHTON, *MORAL VISION: AN INTRODUCTION TO ETHICS* (1988) (same); Smith, *supra* note 92 (same).

96. Or objective conditionals. That is, if you want government to do X, then you want it to follow a particular procedure, since that procedure leads to X.

97. See, e.g., Darwall, et al., *supra* note 95, at 17 ("Most noncognitivists are *expressivists*: they explain moral language as expressing moral judgments, and explain moral judgments as something other than beliefs."); SMITH, *supra* note 92, at 16 ("Expressivists deny that moral judgments represent the world as being one way rather than another. . . .

obliged to follow a particular procedure—whether it is a preference-driven (bargaining) procedure or something else—cannot be true either.

So much for the first, skeptical defense of proceduralism. The several other defenses that I will consider all incorporate a cognitivist rather than skeptical view about the nature of morality. All assume that moral statements can be true or false, and that they are sometimes true; in particular, the remaining defenses assume that moral criteria applicable to regulatory choices truly exist and, most specifically, that moral criteria obliging regulators to follow certain procedures exist. There is nothing self-defeating in this combination of claims (in contrast to the combination involved in skeptical proceduralism). The question, rather, is whether the cognitivist-proceduralist's final claim—the one that makes her theory proceduralist—is a claim we have sufficient reason to accept.

C. *The Foundationalist Defense*

Foundationalism is one possible way to defend proceduralism within a cognitivist moral framework. The idea here is to invoke morality's foundations—to argue that moral truths, at bottom, are truths that issue from certain sorts of procedures—and then to move from this premise to the conclusion that the moral criteria applicable to regulators include procedural criteria. Certainly there are respectable foundationalist views that see moral truths as issuing from certain sorts of collective procedures. By “foundationalism,” I mean a general view about the truth-conditions of moral statements—a general view about why an utterance asserting the existence of, for example, a moral obligation, permission, or criterion is correct.⁹⁸ John Rawls is a famous example of a contemporary philosopher whose foundationalist view involves a collective procedure. Rawls' *A Theory of Justice*⁹⁹ argues that the requirements of justice are those provisions that self-interested parties, bargaining about a social contract, but operating under a “veil of ignorance,” would agree to.¹⁰⁰ More recently, Tim Scanlon has built a moral theory upon a substructure that is more consensualist than the foundational bargain described by *A Theory of Justice*.¹⁰¹ Scanlon suggests that the hypothetical parties whose deliberations result in moral principles should be seen as adopting an impartial, not a self-interested, point of view, and that

[They] rather serve to express the judge's attitudes of approval or disapproval, or perhaps their more complicated dispositions to have such attitudes.” (citations omitted).

98. See generally SHELLY KAGAN, *NORMATIVE ETHICS* 189-303 (1998) (describing and assessing a variety of plausible foundationalist views).

99. JOHN RAWLS, *A THEORY OF JUSTICE* (1971).

100. *Id.* at 118-92.

101. See T.M. SCANLON, *WHAT WE OWE TO EACH OTHER* (1998).

moral principles should (in effect) be defined as principles to which such parties would universally consent.¹⁰²

Does Rawlsian foundationalism, Scanlonian foundationalism, Habermasian foundationalism,¹⁰³ or some other foundationalist view that (like these) envisions a collective procedure as the constitutive grounds for moral truths¹⁰⁴ lead to proceduralism about regulation and governance? Certainly some proceduralists seem to have thought so. For example, there is a foundationalist flavor in much of the literature on civic republicanism. The idea seems to be that impartiality (civic virtue) and unanimity (consensus) are concepts that anyone thinking normatively about governmental choice would need to invoke—that, foundationally, the “public good” is just what public-spirited citizens deliberating to consensus under ideal conditions would agree to—and thus that impartiality, consensus, and explicit consideration of the public good are procedural desiderata for actual governmental choice.¹⁰⁵ One could construct parallel foundationalist stories in defense of pluralism, collaborative-governance theory, or other proceduralist views. As I just suggested, the moral foundations described by Rawls in *A Theory of Justice*¹⁰⁶ would seem more amenable to pluralism than civic republicanism; the Rawlsian social contractors are bargaining, not deliberating (as the very term “social contract” suggests). The pluralist might try to reason from these Rawlsian beginnings to the conclusion that regulatory or governmental outcomes should result from a process of bargaining and compromise among selfish, not public-spirited, actors. So foundationalism is, in principle, a generic strategy for defending proceduralism. But is it a successful strategy?

The answer is no, and the reason for that answer will become clearer if we distinguish between two versions of foundationalism: hypothetical-choice and actual-choice foundationalism. Hypothetical-choice foundationalism says that moral criteria are those criteria that would issue from a certain, hypothetical procedure. The proce-

102. See *id.* at 189-247.

103. See JÜRGEN HABERMAS, *THE INCLUSION OF THE OTHER: STUDIES IN POLITICAL THEORY* 49-101 (Ciaran Cronin & Pablo De Greiff eds., 1998).

104. See KAGAN, *supra* note 98, at 240-56 (discussing “contractarian” foundationalist views, including Rawls’ and Scanlon’s).

105. See, e.g., Seidenfeld, *supra* note 4, at 1528-33. Professor Seidenfeld states: Civic republicanism . . . posits that no individual acting in her political capacity should be subservient to other political actors. Hence, the theory does not equate the public good . . . with majority rule. Social consensus about what is best for the community as a community, not as the aggregation of individuals’ private interests, is the defining feature of the common good. Government’s political decisions—that is, the law—must embody this consensus of the common good.

Id. at 1528-29 (citing Sunstein, *Republican Revival*, *supra* note 5, at 1550, 1548).

106. RAWLS, *supra* note 99.

cedure can be specified as you like: it can be a procedure of impartial deliberation, a procedure of self-interested bargaining, or some other procedure. Actual-choice foundationalism says that moral criteria are those criteria that have actually issued from a procedure of a specified kind that has actually occurred (again, specified as an actual deliberative procedure, an actual bargaining procedure, or some other actual procedure).

Hypothetical-choice foundationalism is a plausible view about morality's foundations; but it does not lead to proceduralism. Rawls and Scanlon are both hypothetical-choice foundationalists. Rawls identifies moral principles as those principles that *would* be chosen by self-interested parties behind a veil of ignorance;¹⁰⁷ Scanlon identifies moral principles as those principles that *would* be chosen by impartial parties deliberating to consensus.¹⁰⁸ For both Rawls and Scanlon, the proposition that one regulatory choice (O_1) is better than another regulatory choice (O_2) flows from the proposition:

(1) O_1 would be chosen over O_2 in a certain, hypothetical, collective procedure.

And this proposition is in turn consistent with the following propositions:

(2) The actual procedure by which O_1 was chosen over O_2 does not conform to the requirements of the hypothetical foundational procedure; indeed

(3) No procedure that has ever occurred conforms to the hypothetical, foundational procedure.

Consider Rawlsian morality, which consists of the following two principles:

Each person is to have an equal right to the most extensive total system of equal basic liberties compatible with a similar system of liberty for all. . . . Social and economic inequalities are to be arranged so that they are both: (a) to the greatest benefit of the least advantaged, consistent with the just savings principle, and (b) attached to offices and positions open to all under conditions of fair equality of opportunity. . . . [These] principles of justice are to be ranked in lexical order . . . [and the] second principle of justice is lexically prior to the principle of efficiency . . . [and] fair opportunity is prior to the difference principle.¹⁰⁹

Then O_1 is better than O_2 , within a Rawlsian view, if O_1 is better in light of the two Rawlsian principles. Equivalently, O_1 would be chosen over O_2 by self-interested parties bargaining behind a veil of ig-

107. See RAWLS, *supra* note 99, at 46-53.

108. See SCANLON, *supra* note 101, at 189-247.

109. RAWLS, *supra* note 99, at 302-03.

norance. But this imposes no constraint upon the actual procedure that leads to O_1 over O_2 . A regulatory choice is fully compliant with Rawlsian justice as long as it complies with the two Rawlsian principles, *whatever* the actual procedure that led to the choice. And the point can be readily generalized to all kinds of hypothetical-choice foundationalism: If the hypothetical-choice procedure constitutive of morality leads to a particular set of criteria ($X_1, X_2 \dots X_n$), and regulatory choice O_1 is better than O_2 in light of the totality of $X_1, X_2 \dots X_n$, then O_1 is morally better than O_2 , *whatever* the actual procedure by which O_1 was chosen.

The objection might be raised that the principles $X_1, X_2 \dots X_n$ resulting from a hypothetical-choice foundation will themselves include certain procedural requirements. But it is not clear why this should be the case. The hypothetical parties, formulating the principles of morality, are obviously not constrained to produce the requirement that actual persons actually carry out the foundational procedure—for example, the two Rawlsian principles do not include such a requirement—nor is it clear why they would end up producing procedural requirements other than a requirement that the foundational procedure be implemented. (Note that Rawls' two principles have hardly any procedural content at all; he is centrally and perhaps exclusively concerned with the effect of governmental choice on the allocation of primary goods and the range of individual liberty.¹¹⁰)

Indeed, it might be argued that hypothetical-choice foundationalism would be *less likely* to produce a proceduralist morality—a set of moral criteria including procedural requirements—than other foundationalist views, since the moral work of procedure would already have been done in the hypothetical situation. The parties in that situation would see no further need for it. In any event, the point to be emphasized here is that hypothetical-choice foundationalism will produce a set of moral criteria ($X_1, X_2 \dots X_n$) that may or may not include procedural requirements (depending on the course that the hypothetical process takes), just as nonprocedural foundationalism will produce a set of moral principles ($X_1^*, X_2^* \dots X_n^*$) that may or may not include procedural requirements. It is $X_1, X_2 \dots X_n$ or $X_1^*, X_2^* \dots X_n^*$ that regulatory institutions must comply with, not the hypothetical procedure itself.

Another objection that might be raised is epistemic. It is difficult to know what $X_1, X_2 \dots X_n$ are. Thus, if a decisionmaker (D) is designing some governmental procedure, specifically, if D is a legislator or President setting up a regulatory-agency procedure, then the best course for D is to require that the agency follow a procedure ap-

110. This aspect of Rawls' theory is discussed in Paul J. Weithman, *Contractualist Liberalism and Deliberative Democracy*, 24 PHIL. & PUB. AFF. 314 (1995).

proximating the hypothetical-choice situation. For example, if *D* believes the hypothetical-choice situation to be a Scanlonian scenario of impartial deliberation and consensus among all involved parties, then the best course for *D* is to set up a legal regime that requires agency deliberation and encourages public-spirited involvement by the citizenry.¹¹¹ The strategy here is to combine hypothetical-choice foundationalism with moral uncertainty on the part of the agency-builder, *D*, to argue for proceduralism.

This strategy is tempting but, I think, unsuccessful. For any plausible specification of the hypothetical-choice scenario constitutive of moral principles, actual governmental procedures will approximate that scenario only very imperfectly. Actual agency rulemaking, however public-spirited and consensual, is generally far removed from the idealized scenario of full reasonableness and universal agreement posited by Scanlon; and, of course, interest groups that engage in political bargaining are *not* operating under a Rawlsian veil. So the epistemic claim becomes the following: Even though the feasible agency procedure that most closely approximates the hypothetical-choice scenario is a crude approximation, the reasonable course for the agency-builder *D* (given his uncertainty about what morality requires) is to instruct agencies to employ this crudely approximate procedure.

Maybe this claim is true, but I doubt it. Wouldn't *D* have some evidence as to what morality requires—as to what the correct moral criteria $X_1, X_2 \dots X_n$ really are?¹¹² And wouldn't he then design agency procedures in light of (his best understanding of) $X_1, X_2 \dots X_n$, rather than simply instructing agencies to do their best to implement the foundational, moral procedure? For example, *D* might come to believe that overall well-being is one of the moral criteria $X_1, X_2 \dots X_n$ and that cost-benefit analysis is the agency decision-procedure best justified in light of overall well-being; and it would then be reasonable for *D* to instruct agencies to engage in cost-benefit

111. See *id.* (presenting epistemic case for proceduralism).

112. The substantive moral theory advanced by *A Theory of Justice*, together with the voluminous scholarly responses to that book, is at least some evidence of what self-interested contractors bargaining behind a veil of ignorance would choose. Similarly, the emerging corpus of scholarly work on Scanlon's theory (presented first in a 1982 article, T.M. Scanlon, *Contractualism and Utilitarianism*, in *UTILITARIANISM AND BEYOND* 103 (Amartya Sen & Bernard Williams eds., 1982), and in fuller form in his book *What We Owe to Each Other*, *supra* note 101) is at least some evidence of the principles that impartial and deliberative contractors would agree to.

More generally, whatever methodology *D* has used to determine the nature of the foundational moral procedure will also, presumably, provide him some evidence of what the $X_1, X_2 \dots X_n$ are. For example, if he has determined the nature of that procedure through a combination of internal deliberation and consultation with scholarly and other experts, then internal deliberation plus consultation with scholarly and other experts should give *D* some evidence about the content of the $X_1, X_2 \dots X_n$.

analysis (as one component of their procedures), even though cost-benefit analysis is surely not a component of the foundational, moral procedure.¹¹³

Note that the epistemic argument has a self-defeating aspect. *D* knows that moral principles are those that would be chosen by a certain, hypothetical procedure; but he does not know enough about the content of those principles to do anything but set up a governmental procedure approximating the hypothetical one. Why does *D* have good evidence (independent of governmental choice) about the *foundations* of morality, but no good evidence (independent of governmental choice) about the *content* of morality? Note further that the epistemic argument is really an instrumental, not an intrinsic one, and thus is not an argument for proceduralism as I have defined it. The claim is not that morality includes a procedural criterion X_p , along with substantive criteria $X_1, X_2 \dots X_n$. The claim, rather, is as follows:

- (1) the hypothetical-choice procedure that is the foundation of morality gives rise to some substantive criteria $X_1, X_2 \dots X_n$; but
- (2) because those who design regulatory agency procedures are ignorant as to the precise content of the $X_1, X_2 \dots X_n$, the best course for them, in light of $X_1, X_2 \dots X_n$, is to instruct agencies to implement some approximation of the foundational procedure.

In any event, this particular (instrumental) claim about regulatory agency procedures is unpersuasive.

So much for hypothetical-choice foundationalism. Hypothetical-choice foundationalism is, again, a very plausible view about the nature of moral truth; but it does not lead to proceduralism. Actual-choice foundationalism does lead to proceduralism, but it involves a highly implausible view about the nature of moral truth. The actual-choice foundationalist says that one outcome (O_1) is morally better than another (O_2) if and only if O_1 actually results from the specified sort of procedure (or, a bit more weakly, if and only if O_1 has an appropriate nexus to an actual procedure that meets the foundational specification, for example, where O_1 is better than O_2 in light of standards enacted by a body that meets the foundational specification).

Actual-choice foundationalism is a purely expository device on my part. It is designed to illustrate what would need to be true about a foundational view for there to be a direct link between such a view and proceduralism. Actual-choice foundationalism is *not* a foundational view that moral philosophers have, in fact, defended, and it is

113. On the status of cost-benefit analysis as a welfare-justified decision procedure, see Adler & Posner, *Rethinking Cost-Benefit Analysis*, *supra* note 28, and *infra* Part IV.E.

easy to see why they have not done so. Imagine that no procedure meeting the foundational specifications has actually occurred, or at least no such procedure with a nexus to O_1 and O_2 . Then the upshot is that the O_1 - O_2 choice must be morally neutral: O_1 and O_2 are equally good, or incomparably good, or something like that, such that O_1 is not definitely better than O_2 nor is O_2 definitely better than O_1 . But O_1 could be substantively horrendous, relative to O_2 . It could be a policy that causes massive welfare losses, large-scale inequalities, and serious damage to environmental values.¹¹⁴ To insist that O_1 in such a case remains no worse than O_2 —because no procedure of the specified type has actually selected O_2 over O_1 —is deeply counterintuitive. Overall well-being, equality of welfare, environmental preservation, deontological constraints, and other moral criteria may all derive their force from the fact that they *would* be chosen in a certain, hypothetical situation; but it is not the case that their moral force is suspended until some actual body confirms them.

D. *The Conventionalist Defense*

A third strategy for defending proceduralism, distinct from moral skepticism and moral foundationalism, is the *conventionalist* strategy. Conventionalism (as I use the term here) is the view that morality, or some part thereof, is a function of social norms.¹¹⁵ Conventionalism could be a general view about morality (in which case it becomes a kind of foundationalism); or, less ambitiously, it could be the view that certain moral criteria (for example, criteria of objective welfare-value, which set forth better or worse “ways of life”) are reducible to social norms.¹¹⁶

Social norms have been conceptualized in various ways but, roughly, norms are constituted by common beliefs and behaviors.¹¹⁷ Crucially, social norms are distinct from *legal* norms and procedures.¹¹⁸ This distinction is a central point in the burgeoning litera-

114. There is obviously nothing in the values of overall well-being, equality of welfare, and environmental preservation that entails that an outcome which is seriously worse than another with respect to one of the values, or all taken together, will actually have been rejected by a governmental body or by some other such mechanism for collective choice.

115. See BRINK, *supra* note 52, at 14-36 (defining conventionalism and distinguishing conventionalism from noncognitivism).

116. See, e.g., JOSEPH RAZ, *ETHICS IN THE PUBLIC DOMAIN* 3-5 (1994) (explaining that valuable activities depend on social practices for their availability and, to a degree, even for their existence).

117. See generally Matthew D. Adler, *Expressive Theories of Law: A Skeptical Overview* 148 U. PA. L. REV. 1363, 1467-72 (2000) (describing various conceptualizations of “social norms”).

118. See *id.*

ture on social norms;¹¹⁹ and it amounts to the fatal flaw in the conventionalist defense of proceduralism. *However* social norms are defined, it is not the case that the (assumed) truth of conventionalism will make certain regulatory procedures intrinsically important.

Consider a simple version of conventionalism that defines social norms, and some or all moral criteria, by reference to the evaluative judgments shared by a large majority of the relevant society.¹²⁰ One outcome is better than another, in light of social norms, only if a large majority of the society would judge the first to be better. Imagine further that the only evaluative judgments shared by a large majority of the relevant society with respect to regulatory agencies are substantive judgments. There is not, let us assume, general agreement that regulatory agencies should follow or refrain from following particular procedures. Then in this society regulators have no intrinsic, conventional obligation to follow or refrain from following certain procedures. The regulator's sheer decision to follow a particular procedure will not (without more) be conventionally good or bad. For example, it might be widely agreed in the society that certain wilderness areas are precious. Then if one outcome (O_1) damages the areas relative to a second (O_2), O_1 will be conventionally worse than O_2 regardless of the procedure by which O_1 was adopted. Conversely, if the two outcomes are neutral with respect to their impact on the wilderness areas, the regulator's choice of one or the other will be conven-

119. The legal scholarship includes ROBERT C. ELLICKSON, *ORDER WITHOUT LAW: HOW NEIGHBORS SETTLE DISPUTES* (1991); Lisa Bernstein, *Opting out of the Legal System: Extralegal Contractual Relations in the Diamond Industry*, 21 J. LEGAL STUD. 115 (1992); Robert Cooter, *Normative Failure Theory of Law*, 82 CORNELL L. REV. 947 (1997); Steven A. Hetcher, *Creating Safe Social Norms in a Dangerous World*, 73 S. CAL. L. REV. 1 (1999); William K. Jones, *A Theory of Social Norms*, 1994 U. ILL. L. REV. 545; Jody S. Kraus, *Legal Design and the Evolution of Commercial Norms*, 26 J. LEGAL STUD. 377 (1997); Lawrence Lessig, *The Regulation of Social Meaning*, 62 U. CHI. L. REV. 943 (1995); Richard H. McAdams, *Cooperation and Conflict: The Economics of Group Status Production and Race Discrimination*, 108 HARV. L. REV. 1003 (1995) [hereinafter McAdams, *Cooperation and Conflict*]; Richard H. McAdams, *The Origin, Development, and Regulation of Norms*, 96 MICH. L. REV. 338 (1997) [hereinafter McAdams, *The Origin, Development, and Regulation of Norms*]; Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903 (1996); Symposium, *Law, Economics, & Norms*, 144 U. PA. L. REV. 1643 (1996); and Symposium, *Social Norms, Social Meaning, and the Economic Analysis of Law*, 27 J. LEGAL STUD. 537 (1998). Yet further entries in the legal literature are cited by Hetcher, *supra*, at 2-3 nn.2-4.

120. Cf. Cooter, *supra* note 119 (arguing that the existence of a norm involves the punishment of nonconformers by "norm enforcers" who have "internalized" the norm); McAdams, *The Origin, Development and Regulation of Norms*, *supra* note 119 (arguing that norms involve a preexisting moral consensus among some portion of the population, which becomes widely known and is then informally enforced by the withdrawal of esteem from norm-violators).

tionally neutral regardless of the procedure by which the choice was resolved.¹²¹

It might be objected that this example does not vitiate the link between proceduralism and conventionalism. Rather, what the example shows is that (on the simple version of conventionalism under consideration here) regulators in the United States have intrinsic procedural obligations if, and only if, a large majority of Americans believe that certain regulatory procedures are bad or good. Thus it is only in the special case where there is *no* procedure whose goodness or badness is supported by an evaluative consensus that proceduralism fails. True enough. *If* the goodness or badness of a particular procedure (P^*) is supported by an evaluative consensus, *then* (on this version of conventionalism) regulators have moral reason to follow or refrain from P^* , independent of its outcomes. But, relatedly:

(1) for any particular procedure P' (for example, a procedure of pluralist bargaining, or of civic republican deliberation, or of collaborative governance), regulators will have moral reason to follow or refrain from P' , independent of outcomes, *only* insofar as an evaluative consensus supporting the goodness or badness of P' exists; and thus

(2) no conventional argument for P' that transcends the population's contingent beliefs about procedures is possible.

In short, the proceduralist who presents a conventionalist argument for pluralism, civic republicanism, collaborative governance, or some other P' is properly engaged in descriptive sociology; all she can claim is that citizens happen to share a common set of beliefs about P' , and that P' will remain of intrinsic importance as long as they do. Note, too, that a firm conventionalist argument for P' (civic republicanism, say) as against P'' (pluralism, say) is not really possible, because consensus beliefs may shift from P' to P'' . The proceduralist who wants an argument for her favored procedure, P' , that outlasts the lucky sociological fact that P' also falls under the description, "believed to be good by most Americans," had better look beyond conventionalism.¹²²

121. I am assuming here that O_1 is not conventionally better than O_2 or vice versa, in light of some feature other than its impact on the wilderness areas. Otherwise, O_1 might be conventionally better than O_2 , all things considered, even though O_1 damages the areas.

122. Again, my discussion here has focused on a particular, simple version of conventionalism; but what I say can, I think, be generalized.

E. The Expressivist Defense

A fourth possible strategy for defending proceduralism might be called *expressivism*.¹²³ Law is sometimes described as expressive insofar as it has a value-creating function.¹²⁴ The idea is that legal outcomes can, themselves, be (partly) constitutive of our evaluative and, specifically, moral framework. How might this idea underwrite proceduralism? Consider the following:

An Expressive Case for Proceduralism

If regulators follow procedure P^* , and they end up choosing O_1 over O_2 , then O_1 is morally better than O_2 , either conclusively (strong-form expressivism) or prima facie (weak-form expressivism).

Expressivism circumvents some of the above-noted difficulties in the skeptical, foundationalist, and conventionalist arguments for proceduralism. Expressivism (as here presented) is cognitivist, not skeptical. Unlike actual-choice foundationalism, it does not adduce P^* as a procedure constitutive of all moral criteria, but rather as a procedure that supplements other grounds for moral distinctions. O_3 can be morally better than O_4 , even if the O_3 -over- O_4 choice is not the outcome of P^* —the expressivist allows as much; but where O_1 is chosen over O_2 by means of P^* , then O_1 is (conclusively or at least prima facie) better. In effect, actual-choice foundationalism makes the implementation of the favored procedure a necessary condition for moral distinctions, while expressivism (more plausibly) makes its implementation a merely sufficient condition. Finally, by contrast with conventionalism, expressivism eschews any link to social norms and practices. P^* is claimed to be morally constitutive, quite apart from whether P^* is generally believed to be morally constitutive (or otherwise favored by existing social norms), and quite apart from whether the O_1 that results from P^* is generally believed to be better than the alternative O_2 (or is otherwise favored over O_2 in light of social norms).

123. "Expressivism," as I use the term here, is one kind of cognitivist moral view and thus is quite different from the noncognitivist metaethical view often referred to by moral philosophers as "expressivism." See sources cited *supra* note 97.

124. See, e.g., Pildes & Sunstein, *supra* note 62, at 70 ("[The] expressive or symbolic dimensions of policy are central [because] [p]art of what policy-making does is to define, interpret, and create collective understandings and values."). Professor Sunstein has also articulated this view of the law as value-generating, thus:

The effects of any legal rule can be described in an infinite number of ways. Any particular characterization or accounting of consequences will rest not on some depiction of the brute facts; instead it will be mediated by a set of (often tacit) norms determining how to describe or conceive of consequences. It is possible to see a large part of the expressive function of law in the identification of what consequences count.

Cass R. Sunstein, *On the Expressive Function of Law*, 144 U. PA. L. REV. 2021, 2048 (1996).

Still, expressivism is a problematic strategy for defending proceduralism. One problem is in specifying a procedure (P^*) that is not too idealized to be infeasible, but that is still sufficiently idealized to be constitutive of the moral standing of the outcomes it produces. Consider a civic-republican specification of P^* : P^* is the procedure where regulators sincerely attempt to be impartial, where they deliberate pretty well, where they are not too uninformed, and where some kind of consensus is reached among a large fraction of those interest-group representatives or other members of the public who participate in the regulatory process. Why should P^* constitute O_1 as morally better than O_2 , even *prima facie*? Let's assume that hypothetical-choice foundationalism is true and also that the foundation of morality is some idealized, civic-republican style procedure, roughly along the lines that Scanlon suggests. Call it P^+ . P^* is not the same procedure as P^+ ; if it were, P^* would be infeasible. So O_1 could be chosen by O_2 through P^* *even though* O_2 would be chosen over O_1 through P^+ . To be more concrete, O_1 could be an outcome that causes large welfare losses and widespread environmental degradation, or in which an innocent victim is sacrificed for the greater good, in violation of her deontological rights. But why should the fact that O_1 was chosen over O_2 in a not-too-bad procedure, by civic-republican lights, confer any moral authority on O_1 , when O_2 would be chosen over O_1 in a perfect civic-republican procedure? Conversely, in the case where both P^* and P^+ point in favor of O_1 , we do not need to invoke P^* to explain the betterness of O_1 ; that is fully explained by P^+ , our civic-republican foundation.

The proceduralist could respond to this objection by denying hypothetical-choice foundationalism. To make the response sharpest, assume she posits a foundation for morality that has nothing whatsoever to do with collective choice or approval. Perhaps she posits a traditionalist view of morality. She might claim, "In general, one outcome is better than another if the first is traditionally believed to be better, in the relevant society. But in the special case where O_1 is chosen over O_2 through P^* , O_1 is at least *prima facie* better on those very grounds." Yet if morality is generally a matter of traditional beliefs, not hypothetical procedures, then why *does* P^* have a morally constitutive role? In short, the expressivist faces a dilemma: either she adopts a hypothetical-choice foundationalism that makes some idealized version of P^* the general grounds of morality (in which case we can ask why feasible and nonideal P^* is also morally constitutive), or she does not (in which case we can accuse her of inconsistency in ascribing moral power to P^* while asserting that the general grounds of morality are quite different).

This line of attack on expressivism might be overcome. Perhaps there *is* some feasible governmental procedure P^* such that its out-

comes are guaranteed to be morally good or right, at least *prima facie*. I have not shown definitively that P^* cannot or does not exist. But even if P^* does exist, the kind of proceduralism that follows from its existence is really quite thin. Regulators are not *obliged* to follow P^* , even *prima facie*. If they follow a different (and nonconstitutive) procedure which is instrumentally appropriate in that it produces morally better O_3 as against morally worse O_4 , then the regulators have not done anything wrong, even *prima facie*. (Contrast this with the standard view, to be considered in a moment, that certain, participatory procedures are intrinsically important because they enhance the welfare of participants. If this view is right, then regulators are obliged, at least *prima facie*, to follow the participatory procedures.) Further, it may be the case that regulators are all-things-considered obliged *not* to follow P^* . This possibility is avoided only if P^* has a conclusive rather than merely *prima facie* role in fixing the moral standing of its outcomes; but that strong-expressivist claim raises the plausible objection that no nonideal governmental procedure could *guarantee* moral betterness.

F. *The Interest-Based Defense*

So much for expressivism. The fifth and final strategy for defending proceduralism to be considered here is widespread in the legal literature.¹²⁵ This is the view that certain governmental procedures are intrinsically beneficial for certain persons (paradigmatically, those persons who participate in the procedures). Call this the *interest-based* defense of proceduralism. Expressivism, conventionalism, and foundationalism all, in different ways, attempt to give the favored procedure a constitutive moral role, or to draw a special link between the procedure and something else that has a constitutive moral role (for example, conventional beliefs or a hypothetical proce-

125. The view that certain governmental procedures are intrinsically beneficial for certain persons has been particularly important in the legal literature on procedural due process. See, e.g., Jerry L. Mashaw, *Administrative Due Process: The Quest for a Dignitary Theory*, 61 B.U. L. REV. 885, 886 (1981) ("The unifying thread in this literature [on procedural due process] is the perception that the effects of process on participants, not just the rationality of substantive results, must be considered in judging the legitimacy of public decisionmaking."); Frank I. Michelman, *Formal and Associational Aims in Procedural Due Process*, in DUE PROCESS: NOMOS XVIII 126, 127 (J. Roland Pennock & John W. Chapman eds., 1977) ("[Participatory] procedures seem responsive to demands for *revelation* and *participation*. They attach value to the individual's *being told why* the agent is treating him unfavorably and to his *having a part in the decision*."). As Professor Summers explains:

My principal thesis in this Essay is that a legal process can be good, as a process, in two possible ways, not just one: It can be good not only as a means to good results, but also as a means of implementing or serving process values such as participatory governance, procedural rationality, and humaneness.

Robert S. Summers, *Evaluating and Improving Legal Processes—A Plea for "Process Values"* 60 CORNELL L. REV. 1, 4 (1974).

ture). The interest-based defense is quite different. It prescind from issues of moral truth and constitution. Instead, it says, much more straightforwardly, that a particular item of moral significance, namely human welfare, can be directly affected by the procedures that regulators employ, quite apart from the welfare effect of outcomes.

The interest-based defense of proceduralism certainly possesses intuitive resonance. Imagine a regulatory agency that furnishes no opportunity whatsoever for participation by the public.¹²⁶ No notice of proposed directives is published; no public commentary on these proposals is even received, let alone solicited; and no public statement justifying enacted directives is ever issued. This is true not only when the directives are general rules, but even when they are individualized orders that fix the rights, duties, and other legal positions of named parties. Even in the case of adjudication, the persons who are affected (however directly) by agency decisions are not afforded a hearing, or any component of a normal hearing, except (let us assume) publication of the directive that the agency finally enacts. Surely this participation-free procedure amounts to an intrinsic welfare setback for members of the public, or at least for those who want to participate in the decisions of the insulated agency!

I deny that it does.¹²⁷ Specifically, I deny that it does insofar as participation diminishes the accuracy of the agency's decisionmaking process; and conversely if participation enhances accuracy, the participatory procedure is already justified on instrumental grounds and so the interest-based argument for participation is superfluous. To see this, imagine that the participation-free procedure leads to outcome *O* rather than *O**, that a participatory procedure would lead to outcome *O** rather than *O*, and that *O* is substantively better than *O** (that is, it is better apart from the procedure by which the *O-O** choice is resolved). Then participation is a bad thing (from the point of view of the rest of our moral framework), and the welfare-claims of those who insist upon participation should be discounted. The welfare-claims of the would-be participants are arguably no different, in this sort of case, from the welfare-claims of those who purport to reap an intrinsic benefit from discrimination or governmental action that

126. In short, imagine that the agency provides *none* of the elements of procedural due process identified by Judge Friendly in his classic article, Henry J. Friendly, "Some Kind of Hearing", 123 U. PA. L. REV. 1267 (1975).

127. I do not deny that the participation-free procedure just described would be unconstitutional under existing doctrines interpreting the Due Process Clause; but one can give content to that clause and defend a judicial role in enforcing it without believing that participation or other procedural elements have intrinsic value for welfare. See *Mathews v. Eldridge*, 424 U.S. 319, 334-35 (1976) (setting forth a three-factor framework for determining what process is "due," which arguably envisions procedure as instrumentally, rather than intrinsically, valuable).

violates moral rights simply because they “like” or “want” these injustices.¹²⁸ On the other hand, if the participatory procedure leads to O^* rather than O , and it is O^* that is the substantively better outcome, then the welfare-claims of the would-be participants need not be discounted. However, the claims need not be discounted only because the regulators already have instrumental grounds here for choosing the participatory procedure, as the one that leads to the substantively best outcome O^* .¹²⁹

The point can clearly be generalized beyond participation. Take a given agency choice situation. Imagine that one procedure ($P_{accurate}$) is the instrumentally justified procedure—the procedure that the legislator, President or administrator (D) who designs the agency’s procedures is justified in requiring the agency to follow in light of substantive moral criteria $X_1, X_2 \dots X_n$ —and that a second procedure ($P_{interest}$) is the procedure that supposedly intrinsically benefits the welfare of the citizenry or some subset thereof. One possibility is that, in this situation, $P_{accurate}$ is $P_{interest}$. If so, regulators can have intrinsic grounds for following $P_{interest}$, but this procedure turns out to be the very one that regulators would have been obliged to follow ($P_{accurate}$) if governmental procedures had no intrinsic welfare value. Alternatively, in the given choice situation, $P_{interest}$ might be *different from* $P_{accurate}$.¹³⁰ If so, we are left to wonder how it can be a good thing for human welfare to skew government towards outcomes that are otherwise morally unfavorable. So either interest-based proceduralism makes no difference to what is morally required of regulators; or, if it

128. See, e.g., GEOFFREY SCARRE, UTILITARIANISM 155-62 (1996) (suggesting that *axiomatic preferences should not be incorporated in the calculation of social welfare*); John C. Harsanyi, *Morality and the Theory of Rational Behaviour*, in UTILITARIANISM AND BEYOND 39 (Amartya Sen & Bernard Williams eds., 1982) (same). See generally RAWLS, *supra* note 99, at 395-587 (arguing that a person’s welfare goals, i.e., his conception of his self-interest, can and should be revised to reflect moral requirements).

129. Larry Alexander has made a similar point:

I do not deny that two procedures that are otherwise equal in factfinding accuracy and that do not violate moral side-constraints may differ in their achievement of other values. And if the procedures are otherwise not only equally accurate, but also equally intrusive on privacy, equally costly in resources, and so forth, we should surely choose the procedure that has more positive or fewer negative byproducts. The more important question, however, is whether either accuracy or social resources should be sacrificed to achieve such values as participation.

Larry Alexander, *Are Procedural Rights Derivative Substantive Rights?*, 17 LAW & PHIL. 19, 36 (1998).

130. I have not considered the case of equality or incomparability. It might be that, in a given choice situation, there are a number of procedures that are equally well or incomparably well justified, instrumentally speaking, and that only one is $P_{interest}$. In that special case I would concede that intrinsicist considerations could justify choosing $P_{interest}$. See Alexander, *supra* note 129, at 36.

does, there is a powerful argument that the putative intrinsic interests in regulatory procedure are spurious.¹³¹

The key to my case against interest-based proceduralism is, obviously, my discounting of certain welfare claims: where the procedure of purported intrinsic benefit ($P_{interest}$) differs from the instrumentally justified procedure ($P_{accurate}$), the intrinsic welfare value of $P_{interest}$ is deflated. Is this discounting fair? As I have suggested, there are certainly other cases where we discount someone's claim to be intrinsically benefited by an outcome, institution, or procedure that is (otherwise) morally wrong (for example, a discriminatory or rights-violating outcome). The proceduralist might object that there are other cases where we *do not* discount someone's claim to be intrinsically benefited by an outcome that is (otherwise) morally wrong. For example, if I like fishing and everyone else likes jet-skiing, and a lake can only be used for one or the other, then the outcome in which the lake is used for fishing is morally wrong apart from its effect on my welfare. But that does not mean that my interest in fishing should be discounted; rather, the interest in fishing retains full force and should be thrown into the balance along with the interests of the jet-skiers.¹³² This is a fair objection. I think that procedural interests are more like discountable interests in discrimination or rights violation than they are like nondiscountable interests in fishing; but the point is surely debatable.

My claim that procedural interests are discountable is, in part, an appeal to the reader's intuitions. Where procedure $P_{interest}$ leads to morally worse results than $P_{accurate}$, it seems intuitively problematic to think that some person P 's welfare is genuinely improved by the

131. Might the proceduralist exploit the disjunction between what is ex ante justified in light of $X_1, X_2 \dots X_n$, and what is ex post justified? $P_{accurate}$, as I have characterized it, is the procedure that the designer D is ex ante justified in requiring the agency to follow—in other words, the procedure that D is justified in requiring, given less than perfect information on D 's part. (For more discussion of moral choice under uncertainty, see *infra* Part IV.E. If D had perfect information, he would know already what $X_1, X_2 \dots X_n$ required in every choice situation and presumably could just instruct agencies to do that.) But $P_{accurate}$ might not be ex post justified in a given agency choice situation; it might not end up producing the outcome that is best in light of $X_1, X_2 \dots X_n$.

So it is possible that in a given agency choice situation (1) $P_{accurate}$ deviates from $P_{interest}$, but (2) $P_{interest}$ leads to the outcome that is ex post justified. In such a scenario, should the welfare-value of $P_{interest}$ be discounted? Perhaps not. Even so, this (partial) response to my discounting argument does not change how the designer D should act. He is still ex ante justified in requiring agencies to follow $P_{accurate}$, rather than requiring them to deviate in some cases from $P_{accurate}$ and follow $P_{interest}$ instead. D , after all, does not know which are the cases in which $P_{accurate}$ produces results that are ex post unjustified; if he did know that, he would revise $P_{accurate}$ so as to eliminate the disjunction, and the cases in which procedural interests are nondiscountable because $P_{accurate}$ and ex post-justifiability deviate would disappear.

132. Note the paradox otherwise: if the fishing interest is discounted then, symmetrically, the jet-skiing interest should also be discounted, and we end up with no interests at all.

performance of $P_{interest}$. We *do* have strong intuitions about the sorts of procedures that government ought to follow (for example, about the need for participation by the person who will be directly affected by an adjudicative hearing), but these intuitions (I think) are generally connected to the belief that the procedure at stake enhances, rather than diminishes, accuracy. Relatedly, our intuition that participation and other procedural features have welfare-value is not (I think) severable from the belief that the features are accuracy-enhancing.

Further, I would suggest, the discountability of procedural interests is consistent with the account of well-being I have defended in this Article, namely sophisticated preferentialism. Consider a given person (P) who claims to have an intrinsic welfare interest in a particular procedure, $P_{interest}$, as compared to another procedure, $P_{accurate}$, even in the case where $P_{interest}$ leads to morally worse results. Given sophisticated preferentialism, P truly benefits from $P_{interest}$ only if:

- (1) P prefers $P_{interest}$ to $P_{accurate}$ and
- (2) P would prefer $P_{interest}$ to $P_{accurate}$ under suitably ideal conditions.

The first prong of this schema is unproblematic: people can have all sorts of odd preferences. But why think that P 's preference for $P_{interest}$ would survive idealization? Perhaps P thinks that $P_{interest}$ serves his own welfare: he prefers $P_{interest}$ to $P_{accurate}$ even where the first leads to O^* , the second leads to O , and O^* is morally worse, because O^* is better *for P* than O . Yet in such a case P does not really have an idealized preference for a particular procedure ($P_{interest}$); rather, he has an idealized preference for outcomes (O^* over O) and, derivatively, for whatever procedure leads to them.

What we need to show is that P can have a preference and an idealized preference for $P_{interest}$ which is

- (a) not merely an instrumental preference, derivative of the outcomes that $P_{interest}$ produces (for example, a preference for morally better outcomes, or for outcomes that are beneficial to P himself); and
- (b) which retains force even where $P_{interest}$ produces O^* rather than O and O^* is morally worse than O .

I am skeptical that such idealized preferences actually obtain; and if they do obtain I would guess that they occur so infrequently and (relatedly) are sufficiently idiosyncratic, that they would make little difference to the moral calculus of regulators. Given a case where $P_{accurate}$ leads to O and $P_{interest}$ leads to O^* , there would need to be sufficient persons who prefer and ideally prefer $P_{interest}$ (notwithstanding its moral costs), such that regulators are all things considered obliged to use $P_{interest}$ —that is, such that the welfare benefit of $P_{interest}$

compared to $P_{accurate}$ outweighs the welfare and other moral costs of O^* compared to O . My conjecture (debatable, to be sure) is that the frequency of such cases is trivial.

IV. A WELFARIST THEORY OF REGULATION

This Part defends a welfarist theory of regulation and describes the implications of that kind of theory. What is a welfarist theory? By that I mean a normative theory of regulation which asserts the following propositions:

- ??Overall well-being is morally relevant to regulatory choice, if not morally conclusive of regulatory choice. It is one criterion (perhaps supplemented by others) that bears upon the moral status of regulation.
- ??The right theory of well-being is sophisticated preferentialism.
- ??Other moral criteria possibly bearing upon regulatory choice include distributive criteria, deontological criteria, and perfectionist criteria, *but not* the criterion of Kaldor-Hicks efficiency, nor procedural criteria.

Welfarism, thus described, is really a family of moral theories rather than a unique one. For example, the version of utilitarianism that incorporates a sophisticated preferentialist view of well-being is a welfarist theory; so is a consequentialist theory that takes both overall well-being (again understood in sophisticated preferentialist terms) and the equal distribution of well-being as criteria for ranking consequences; and so is a mixed theory that combines the consequentialist theory just mentioned with deontological constraints such as the constraint against intentional harming. It would be nice to be in a position to argue for a unique normative theory of regulation. That would be a more impressive and practically significant accomplishment than defending a family of views, such as welfarism. But I am in no such position now, nor am I likely ever to be. On the other hand, welfarism excludes the theories that have been most important within American legal scholarship, namely neoclassicism and proceduralism—no theory can be both a welfarist theory and a neoclassical theory, or both a welfarist theory and a proceduralist theory—and thus a defense of welfarism should, by itself, have real payoffs.

In the initial parts of this Article, I argued for certain components of welfarism. Specifically, I tried to show that sophisticated preferentialism is indeed the correct account of well-being; that Kaldor-Hicks efficiency is not a moral criterion; and that no procedural standard (for example, a standard requiring regulators to allow citizen participation or to engage in deliberation prior to choosing outcomes) is a moral criterion either. But the central component of welfarism—the moral relevance of overall well-being—remains to be shown. Is over-

all well-being a determinate concept? Given two regulatory choices, such that some persons gain by the first choice and others gain by the second, is it meaningful to say that one choice or the other is better in light of overall well-being? Does not such a statement implausibly presuppose that interpersonal comparisons are possible? And even if such comparisons *are* possible, why does overall well-being have moral force? Why should an increment to P_2 's well-being be grounds for P_1 to accept a decrement to her own, just because P_2 's increment is (in some sense) greater than P_1 's decrement?

Section A argues that overall well-being is a determinate concept. Section B argues that the concept is not only determinate, but possesses moral force.¹³³ Section C discusses the possible role of distributive, deontological, or perfectionist criteria within a welfarist theory. Sections D and E describe the implications of welfarism for regulatory choices and institutions. Section D compares welfarism to neoclassicism. Although welfarism is, in effect, refurbished neoclassicism, there remain substantial differences between the two theories, and I elaborate these. Section E compares welfarism to proceduralism. Here, the implications of the new theory are even larger, since for the welfarist the *only* moral significance of regulatory procedures is in producing welfare-maximizing choices and, perhaps, choices that also fairly distribute welfare, maximize perfections, and comply with whatever deontological norms limit government regulators.

A. *Interpersonal Comparisons and the Determinacy of Overall Well-Being*

Sophisticated preferentialism makes welfare a combination of (restricted) actual preference and (restricted) ideal preference. It says: P would be better off with S_1 , as compared to S_2 , just in case:

- (1) P prefers S_1 over S_2 ; and
- (2) P would prefer S_1 over S_2 under ideal conditions; and
- (3) P 's actual preference and ideal preference are both suitably restricted.

The difference between sophisticated preferentialism and the simple preference-based view of welfare, which is the view that is conventional within welfare economics and which neoclassicism incorporates, lies in conditions (2) and (3). The function of condition (2) is to add ideal preference; the function of condition (3) is to restrict both actual and ideal preference. But it does not seem that either of these conditions will help sophisticated preferentialism solve the problem

133. Sections A and B are a more elaborate and, in some respects, materially different version of the argument presented in Adler & Posner, *supra* note 28, at 197-209.

of interpersonal comparisons—a problem that, famously, bedevils the simple view.¹³⁴

Why are interpersonal comparisons a serious problem for the simple preference-based view of welfare? That view, again, says that P would be better off with S_1 , as compared to S_2 , just in case P prefers S_1 to S_2 . The problem arises once we move from Pareto-superiority to Pareto-noncomparability, that is, once we add a second person (Q) to the picture, such that P is benefited by S_1 over S_2 but Q is benefited by S_2 over S_1 . How should we decide whether S_1 or S_2 is better for overall well-being—that is, better in light of the aggregate welfare of P and Q ?

One approach would be to rely upon P 's and Q 's ordinal ranking of the outcomes, but this approach seems morally arbitrary and no one (as far as I am aware) has seriously advocated it. Imagine that there are N possible outcomes. P and Q assign numbers of 1 to N to each outcome, depending on their preferences, with N given to the very best outcome and 1 given to the very worst.¹³⁵ In particular, P gives a ranking of P_1 to S_1 , and P_2 to S_2 (where $P_1 > P_2$); while Q gives a ranking of Q_1 to S_1 and Q_2 to S_2 (where $Q_1 < Q_2$). Then the ordinalist would say that S_1 is better for the overall welfare of P and Q if and only if $P_1 - P_2 > Q_2 - Q_1$.

The ordinalist approach to interpersonal comparisons is quite problematic, for several reasons. First, it assumes that the N outcomes are equally "spaced" in light of overall well-being—that the effect on overall well-being of moving P from her lowest-ranked outcome (the outcome assigned number 1) to her next-lowest-ranked outcome (the outcome assigned number 2) is the same as the effect on overall well-being of moving her from her next-lowest-ranked outcome to the outcome assigned number 3. More generally, it assumes that the effect on overall well-being of moving P from an outcome assigned number i to an outcome assigned number $i+k$ is the same regardless of where i and $i+k$ are in P 's rankings.

Second, it assumes that different persons are equally productive of well-being. Any policy that moves P up r units in his ranking, while moving Q down q units will necessarily increase overall well-being, so long as q is smaller than r , regardless of what the set of N

134. See generally KEN BINMORE, *GAME THEORY AND THE SOCIAL CONTRACT: PLAYING FAIR* 282-96 (1994) (presenting problem of interpersonal welfare comparisons and various possible solutions); Peter J. Hammond, *Interpersonal Comparisons of Utility: Why and How They Are and Should Be Made*, in *INTERPERSONAL COMPARISONS OF WELL-BEING* 200 (Jon Elster & John E. Roemer eds., 1991) (same); Daniel M. Hausman, *The Impossibility of Interpersonal Utility Comparisons*, 104 *MIND* 473 (1995) (same).

135. This assumes that neither person is indifferent between any two outcomes; otherwise, the ordinalist method here described would need to be modified.

possible outcomes is. For example, suppose we are in a universe with only three outcomes: S_1 (complete bliss for P , torturous pain for Q), S_2 (complete bliss for P , dampened only by a headache and a hang-nail; mild pleasure for Q), and S_3 (complete bliss for P , dampened only by a headache; complete bliss for Q). P will rank S_1 , S_2 , and S_3 3, 1, and 2, respectively, while Q will rank them 1, 2, and 3. Thus, the ordinalist approach ranks S_1 better than S_2 in light of overall well-being, even though ever-blissful P 's relief from headache and a hang-nail is purchased at the cost of moving Q from mild pleasure to torturous pain.

A somewhat more attractive approach, advocated by J.R. Isbell¹³⁶ and others, is to determine overall well-being based upon P 's and Q 's *cardinal* rankings of the outcomes.¹³⁷ These cardinal numbers will be derived from each person's ordinal rankings of the various outcomes, plus her rankings for the various outcomes as compared to risky lotteries of the outcomes. Roughly, P will be told to assign the number 1 to his very best outcome, 0 to his very worst outcome, and a number r to every intermediate outcome such that he would be indifferent between that intermediate outcome and a lottery with a probability of r for the very best outcome and $1-r$ for the very worst. Q will be told the same. Overall well-being is then determined, quite simply, by adding P 's numerical increment or decrement (in terms of her cardinal numbers) to Q 's increment or decrement (in terms of her cardinal numbers). If, for example, P assigns $P1$ to S_1 and $P2$ to S_2 (where $P1 > P2$), and Q assigns $Q1$ to S_1 and $Q2$ to S_2 (where $Q1 < Q2$), then S_1 is better overall as compared to S_2 if and only if $P1-P2 > Q2-Q1$. This approach is more attractive than the ordinalist approach because it does *not* assume equal "spacing." The numerical difference between the number assigned by P to his worst outcome (0) and his next-to-worst outcome need not be the same as the difference between the number assigned to his next-to-worst outcome and his third-worst outcome.

However, there turn out to be various technical objections to the cardinalist approach.¹³⁸ Quite apart from these objections, the approach is suspect because—like the ordinalist approach—it assumes that different persons are equally productive of well-being. Moving P

136. See Hammond, *supra* note 134, at 215-16 (discussing Isbell's construct); Hausman, *supra* note 134, at 479-82 (same); Ruth Weintraub, *Do Utility Comparisons Pose a Problem?*, 92 PHIL. STUD. 307, 317-18 (1998) (same).

137. Isbell's technique has, in effect, been adopted by the so-called QUALY approach to evaluating governmental projects that affect longevity or health. See Robert Fabian, *The Qaly Approach*, in VALUING HEALTH FOR POLICY: AN ECONOMIC APPROACH 118, 119, 124 (George Tolley et al., eds., 1994) (stating that "QUALYs [i.e., quality-adjusted life years] returning to different individuals should be weighted equally" and that "a full healthy life for each individual carries the same weight" (internal quotations omitted)).

138. See Hausman, *supra* note 134, at 479-82.

from his very worst outcome to his very best outcome is assumed to have the same effect on overall well-being as moving Q from *her* very worst outcome to her very best outcome; more generally, moving P a fraction (r) of the way from his very worst outcome to his very best outcome is assumed to have the same effect on overall well-being as moving Q the very same fraction (r) of the way from *her* very worst outcome to *her* very best outcome. This effect is reflected in the cardinalist's assignment of the same numbers, 0 and 1, to each person's worst and best outcomes, and in the fact that the numerical increments or decrements summed across persons to determine the effect of a given move (S_2 to S_1) on overall well-being are unweighted. It is one thing to say that persons have an equal *claim* to well-being, that is, that equality of well-being or other such distributive criteria are relevant to regulatory choice. It is quite another to say that overall well-being is *itself* inherently distributive, in that whatever the set of possible outcomes, overall well-being is a function of P 's and Q 's cardinal rankings normalized on a zero-one scale. Imagine that P is an ascetic, while Q 's welfare is exquisitely sensitive to small changes in the world that would barely make a difference to P . Then, for a given cardinal increment, that increment, experienced by P , should have less significance for aggregate well-being than if experienced by Q . (This is just what it means, after all, to say that P is an ascetic, while Q is not.) But the cardinal approach makes this impossible; it ignores affective differences and other such differences that bear upon the capacity of persons to contribute to aggregate welfare.¹³⁹

A third approach to interpersonal comparisons of welfare, on a simple preference-based view of welfare, is the well-known suggestion of Harsanyi¹⁴⁰ and others¹⁴¹ that we look to the *extended preferences* of an impartial observer. P prefers S_1 to S_2 , while Q prefers S_2 to S_1 . Observer I is asked to

- (i) imagine herself in P 's shoes, with all of P 's characteristics including P 's preferences, moving from S_2 to S_1 ;
- (ii) imagine herself in Q 's shoes, with all of Q 's characteristics including Q 's preferences, moving from S_1 to S_2 ; and

139. See AMARTYA K. SEN, COLLECTIVE CHOICE AND SOCIAL WELFARE 98 (1970) (criticizing the Isbell approach because it fails to allow "interpersonal variability of capacity for satisfaction"); RAWLS, *supra* note 99, at 323 (making a similar criticism); GRIFFIN, *supra* note 73, at 120 (making a similar criticism and noting "It is not the case that we all reach the same peaks and valleys.").

140. See Harsanyi, *supra* note 128; see also John A. Weymark, *A Reconsideration of the Harsanyi-Sen Debate on Utilitarianism*, in INTERPERSONAL COMPARISONS OF WELL-BEING, *supra* note 134, at 255, 289-97 (presenting Harsanyi's model).

141. See, e.g., James Griffin, *Against the Taste Model*, in INTERPERSONAL COMPARISONS OF WELL-BEING, *supra* note 134, at 45, 52 n.15 (citing other scholars who develop variants of Harsanyi's model, including Kenneth Arrow, Amartya Sen, R.M. Hare, and Donald Davidson).

(iii) decide whether she would prefer the first move or the second.

The extended-preference approach is inviting first because it is sensitive to the possibility of affective differences and the like that make persons differentially welfare-productive, and second because it is open to the possibility that outcomes might be differentially "spaced" in terms of persons' welfare. All this would bear upon I 's ranking of the S_2 - S_1 move for P , as compared to the S_1 - S_2 move for Q . The problem with the extended-preference approach lies elsewhere. It is that different impartial observers might have different rankings of the relevant items. For example, observer I might prefer being P and moving from S_2 to S_1 , over being Q and moving from S_1 to S_2 ; while another observer (I^*) might have just the opposite preference.¹⁴²

What is Harsanyi to say to this objection? He might say that, if I and I^* were sufficiently idealized, that is, if their preferences were sufficiently informed, sufficiently deliberate, and so on, then both would prefer either the first move to the second or the second move to the first. But is this true? One person could have an idealized preference for vanilla ice cream, while another could have an idealized preference for chocolate ice cream, just by virtue of differences in their physical makeups and backgrounds. If this is true, then it is not clear why idealized I and idealized I^* could not have different preferences over the moves for P and Q . Observers I and I^* may just have different "tastes" with respect to the prospect of becoming one person or, instead, a different person (P or Q) in a world that undergoes a particular change (the move between S_1 and S_2).

A final approach to determining overall well-being, consistent with the simple preference-based view, is to translate each person's preferences over outcomes into cardinal units of one or another basic resource for well-being, where an objective, public, and cardinal scale for measuring such units exists. For example, an objective, public, and cardinal scale of dollars exists; it is an objective fact that P has, say, \$100, while Q has \$80, and that moving P to a level of \$90 would decrease her stock of dollars more than moving Q to a level of \$85 would increase Q 's. So where the outcomes to be compared are S_1 and S_2 , we can give P a dollar equivalent D_p for the S_2 -to- S_1 move, and Q a dollar equivalent D_q for that move. We can then subtract the Losers' equivalents from the Winners' equivalents to determine the effect of the move on overall well-being. In short, *cost-benefit analysis* in terms of dollars, or liberty, or any other objectively-scaled resource could be understood as a methodology for interpersonal comparison—

142. See GRIFFIN, *supra* note 73, at 53-54 (articulating this kind of criticism of Harsanyi); Hausman, *supra* note 134, at 477-78 (same); see also BINMORE, *supra* note 134; HURLEY, *supra* note 67, at 103-11; Thomas M. Scanlon, *The Moral Basis of Interpersonal Comparisons*, in INTERPERSONAL COMPARISONS OF WELL-BEING, *supra* note 134, at 17, 22-38.

where the existence of an objective scale is meant to get around the unwelcome effect of subjectivity (the divergence between the α -extended preferences of I and I^*) that afflicted Harsanyi's approach, while the reliance on a welfare resource is meant to make plausible the claim that cost-benefit analysis and overall well-being are just equivalent. But are they? Can they be, even where cost-benefit analysis is performed in terms of a basic resource—a "primary good,"¹⁴³ to quote Rawls, such as dollars?

It seems not. The central claim of cost-benefit analysis, again, is that the move from S_2 to S_1 is better in light of the overall well-being of P and Q just in case the winner P 's resource equivalent R_p (for example, a dollar equivalent) is larger than the loser Q 's resource equivalent R_q . This claim is problematic, regardless of how R_p and R_q are defined.¹⁴⁴ If R_p and R_q are defined as the amounts that P and Q would be willing to pay or accept for the S_2 -to- S_1 move, given their actual holdings of the basic resource, then the claimed identity between cost-benefit analysis and overall well-being bumps up against the objection that these numbers can be distorted by the differential size of P 's and Q 's holdings. Imagine that S_1 is an outcome in which a particulate fouls the sky, thereby improving the beauty of the sunsets that P views, but causing respiratory distress to Q . Because P owns tremendous amounts of the resource, he is willing to spend a very large amount (R_p) to purchase the sunset; that large expenditure will still leave P free to purchase everything else he needs. Because Q is relatively poor in her resource holdings, she is willing to spend less (R_q) to purchase relief from respiratory distress,¹⁴⁵ additional amounts beyond R_q are needed by Q for goods even more important than relief from respiratory distress, for example, food or shelter. So R_p is greater than R_q but the S_2 -to- S_1 move still seems to decrease, not increase, overall well-being; respiratory distress is more important for welfare than a pretty sunset. To put the point more generally and abstractly, the proposal to define R_p and R_q in terms of P 's and Q 's actual holdings of the resource ignores the possibility that the resource may have diminishing marginal utility; and

143. RAWLS, *supra* note 99.

144. See Adler, *supra* note 63, at 1373-74; Adler & Posner, *supra* note 28, at 191-95.

145. If R_p is the amount that P is willing to spend in S_1 to purchase the sunset then, technically, R_q is the amount that Q would be willing to accept in S_1 in compensation for the respiratory distress—not the amount that Q would be willing to pay (in S_2) to purchase relief from distress. However, it is particularly easy to see that the amount P is willing to spend to purchase the sunset could be greater than the amount Q is willing to spend to purchase relief, by virtue of differences in their holdings of the basic resource. Further, absent special circumstances (such as would arise if the respiratory distress and sunset made a large difference to Q 's and P 's overall well-being), the amount that Q is willing to accept in compensation for the respiratory distress should be equal to the amount that she is willing to spend to purchase relief from it.

yet it is widely recognized that the marginal utility of all goods, including basic resources (like dollars), can and usually does diminish.

One way around the diminishing-marginal-utility objection to cost-benefit analysis, perhaps, is to define R_p and R_q in terms of P 's and Q 's *normalized* resource-holdings rather than their actual holdings. If P and Q had equal amounts of the resource, how much would P pay for the S_2 -to- S_1 move, as compared to the amount that Q would require for the S_2 -to- S_1 move? The difficulty here is the same as with Isbell's cardinal-utility approach to interpersonal comparisons, and with the ordinalist approach: Persons can be differentially productive of well-being. Although P and Q may have an equal distributive claim to the resource, that does not entail, and it does not seem plausible to say, that P 's (diminishing) function relating basic resources and well-being is necessarily the same as Q 's (diminishing) function. If P is ascetic while Q is super-sensitive, P 's R_p (defined in the "normalized" way described above) can be bigger than Q 's R_q even though the S_2 -to- S_1 move, like all other changes in the world, makes barely a difference to P 's welfare.¹⁴⁶

To sum up, interpersonal comparisons of well-being are seriously problematic for the simple view that equates well-being with preference satisfaction. All of the preference-based approaches to interpersonal comparison—the ordinal approach, Isbell's cardinal approach, Harsanyi's extended-preference approach, and the cost-benefit approach—encounter cogent objections. And these problems with interpersonal comparisons would seemingly carry over from the simple preference-based view to the view espoused by welfarism, namely sophisticated preferentialism.

The sophisticated preferentialist could follow an ordinal, cardinal, extended-preference, or cost-benefit approach, defined in terms of actual preference. That is, the sophisticated preferentialist could say:

?? P is better with S_1 as opposed to S_2 because P 's actual restricted preference favors S_1 and his ideal restricted preference favors S_1

?? Q is better with S_2 as opposed to S_1 because Q 's actual restricted preference favors S_2 and his ideal restricted preference favors S_2

?? Whether S_1 or S_2 is better for the overall welfare of P and Q depends on P 's and Q 's actual restricted preference; in particular the [ordinal approach/cardinal approach/extended-preference approach/cost-benefit approach] is the right way to integrate ac-

146. The point that money can have differential welfare productivity as between equally wealthy persons, e.g., because of their physical differences, has been recognized in the literature on pain-and-suffering damages. See W. KIP VISCUSI, FATAL TRADEOFFS: PUBLIC AND PRIVATE RESPONSIBILITIES FOR RISK 70-73 (1992); Alan Schwartz, *Proposals for Products Liability Reform: A Theoretical Synthesis* 97 YALE L.J. 353, 364 (1988). See generally Steven P. Croley & Jon D. Hanson, *The Nonpecuniary Costs of Accidents: Pain-and-Suffering Damages in Tort Law*, 108 HARV. L. REV. 1785 (1995).

tual preferences so as to arrive at a judgment of overall well-being.

But then the above-described objections to the ordinal, cardinal, extended-preference, and cost-benefit approaches apply directly to the sophisticated preferentialist's proposal.

Alternatively, the sophisticated preferentialist could use *P*'s and *Q*'s *ideal* preferences to determine overall well-being by plugging ideal preferences into the ordinal, cardinal, extended-preference, or cost-benefit approaches; but it is not clear why the difficulties with such approaches are avoided by the shift to ideal preference. Take Isbell's cardinal approach, in which the idea would be to assign 1 to the outcome ranked very best by *P*'s ideal preference, 0 to the outcome ranked very worst, and a number *r* to each intermediate outcome such that *P*, choosing on the basis of his ideal preferences, would be indifferent between that outcome and an (*r*, 1-*r*) lottery of the very best and worst outcomes. This approach still assumes, implausibly, that moving *P* from his very worst to his very best outcome must have exactly the same effect on overall well-being as moving *Q* from her very worst to her very best outcome.

What to do? One option is to abandon the project of interpersonal comparisons. The sophisticated preferentialist could say that, where two outcomes are Pareto-noncomparable, overall well-being is simply indeterminate. But this option is deeply unattractive. As the philosopher Daniel Hausman has rightly commented: "[I]f a conception of well-being does not permit one to make interpersonal comparisons in an acceptable way, then that conception of well-being is itself unacceptable."¹⁴⁷ It may be true that *some*, even most Pareto-noncomparable outcomes are welfare-noncomparable; but the view that all such outcomes are welfare-noncomparable is highly counter-intuitive.

Imagine that *S*₁ involves a painful death for thousands (who prefer not to die); in *S*₂, the thousands do not die, but *P* suffers a mild headache (which he prefers not to have). *S*₁ and *S*₂ are Pareto-noncomparable, but surely we would want to, and can, say that *S*₂ is better for overall well-being than *S*₁. If sophisticated preferentialism fails to warrant such an obviously true statement, then that should be taken as a strong and maybe decisive objection against this welfare view.

Note further that if sophisticated preferentialism disables judgments of overall well-being, then it also disables the comparisons of welfare levels necessary for distributive criteria (for example, the cri-

147. Hausman, *supra* note 134, at 474. Hausman continues: "[I]nterpersonal comparisons are an ineliminable part of human life." *Id.* at 489. For similar statements, see Haranyi, *supra* note 128, at 49; Weintraub, *supra* note 136, at 307.

terion that seeks to equalize well-being, or to maximize the welfare position of the least well-off).¹⁴⁸ To put the point epistemologically, in choosing between sophisticated preferentialism and other welfare views (for example, hedonism, simple objectivism, or sophisticated objectivism), we ought to strive for a “reflective equilibrium”¹⁴⁹ where the chosen view is the one best supported by intuitive judgments as well as systematic considerations. Sophisticated preferentialism, if it entailed the welfare-noncomparability of Pareto-noncomparable states and the impossibility of welfare-level comparisons, would be so counterintuitive and so systematically problematic that it could not be the equilibrium welfare view.

A more palatable option for the sophisticated preferentialist is to use objective values as a basis for interpersonal comparisons. James Griffin (one of the leading philosophers of well-being, and an advocate of something like the sophisticated preferentialist view) has argued that judgments of overall well-being as between Pareto-noncomparable states reduce to judgments of overall objective value.

We have a picture of normal human desires: virtually all persons, when informed, want to live autonomously, to have deep personal relations, to accomplish something with their lives, to enjoy themselves. With experience, we build up such a profile of the components of a valuable life, including their relative importance—a chart to the various peaks that human life can reach. These values, if our profile is complete, cover the whole domain of prudential [i.e., welfare-related] value. They are valuable in any life; individual differences matter not to what appears in this profile of general prudential values, but to how, or how much, a particular person can realize one or other particular value.

. . . [Consider] Mill’s interpersonal comparison of Socrates and the Fool. . . . What [the person making the comparison] needs to make is a judgment of a very different sort from what we ordinarily understand by a personal preference. He needs to know how much persons generally, when informed, would want each life, how desirable they are. This judgment *can* be expressed as a personal preference, but the nature of the judgment is very special: it is a judgment about prudential values that is independent of what any particular individual’s desires or preferences happen to be.¹⁵⁰

The objectivist approach to interpersonal comparisons is sometimes framed (as Griffin seems to frame it) in terms of a list of objec-

148. See BOADWAY & BRUCE, *supra* note 29, at 137-70 (discussing extent to which different “social welfare” criteria, such as the utilitarian criterion or distributive criteria, demand interpersonal comparability of utility levels or differences).

149. See NORMAN DANIELS, *JUSTICE AND JUSTIFICATION: REFLECTIVE EQUILIBRIUM IN THEORY AND PRACTICE* (1996) (presenting and extending Rawlsian idea of “reflective equilibrium”).

150. GRIFFIN, *supra* note 73, at 114-17.

tive values $V_1, V_2 \dots V_n$, such that the comparative effect of S_1 and S_2 on overall well-being is a function of how they compare with respect to each V_i , plus tradeoff rates between different values.¹⁵¹ (So if V_1 is the value of physical pleasure, and V_2 is the value of artistic accomplishment, then we might determine that: (1) moving from S_2 to S_1 increases Q 's physical pleasure by twenty units and decreases P 's physical pleasure by ten units, thus producing an overall increase of ten units of physical pleasure; (2) moving from S_2 to S_1 decreases Q 's artistic accomplishment by eleven units, and increases P 's artistic accomplishment by eight units, thus producing an overall decrease of three units of artistic accomplishment; and (3) one unit of artistic accomplishment is worth four units of physical pleasure, objectively speaking, and so the move from S_2 to S_1 decreases overall objective value.) This may be the right way to make overall judgments of objective value, but it need not be. The objectivist about overall welfare is not committed to the existence of cardinally measurable "values" plus tradeoffs between these. Rather, and more weakly, she is committed to the existence of some sense in which it can be "objectively" true that P is benefited more or less by the move from S_2 to S_1 than Q is harmed.

What is that sense? Objectivity, with respect to matters of value, morality, and norms, is plausibly understood as the convergence of judgments or preferences under ideal conditions. To quote Michael Smith, "[It] is desirable [that is, more valuable, morally better, or otherwise normatively better] that p in C just in case we would all desire that p in C if we were fully rational."¹⁵² As Smith elaborates:

[T]he truth of a normative reason claim requires a convergence in the desires of fully rational agents. However note that the convergence required is not at the level of desires about how each such agent is to organize her own life in her own world. In their own worlds fully rational agents will find themselves in quite different circumstances from each other, circumstances that are conditioned by their different embodiments, talents, environments, and attachments in their respective worlds. Their desires about how to organize their own lives in their own worlds will therefore reflect these differences in their circumstances. The convergence required is rather at the level of their hypothetical desires about what is to be done in the various circumstances in which they might find themselves.¹⁵³

Some person's performance of action A is objectively, normatively better than an alternative, just in case everyone, under ideal condi-

151. See also Scanlon, *supra* note 142, at 39-44 (apparently advocating this approach to interpersonal comparisons).

152. SMITH, *supra* note 92, at 166.

153. *Id.* at 173.

tions, would prefer in this situation to perform A , or so Smith plausibly proposes.¹⁵⁴ In particular, I would argue, S_1 is objectively better for P 's welfare than S_2 just in case everyone under ideal conditions would (restrictedly) prefer to be P experiencing S_1 rather than P experiencing S_2 . And the S_2 -to- S_1 move is objectively better for the overall well-being of P and Q —where P benefits from the move and Q loses—just in case everyone under ideal conditions would (restrictedly) prefer to be P moving from S_2 to S_1 rather than Q moving from S_1 to S_2 .¹⁵⁵ For example, where S_2 involves complete bliss for one person but torturous pain for a second, and S_1 involves bliss dampened by a headache and hangnail for the first person and mild pleasure for the second person, S_1 would represent an improvement in overall well-being if and only if everyone under ideal conditions would prefer (restrictedly) to replace torturous pain with mild pleasure (to be the second person moving from S_2 to S_1 rather than to be rid of a headache and hangnail that dampen complete bliss (to be the first person moving from S_1 to S_2)).

Note that the objectivist model of interpersonal comparisons, as I have framed it, is really quite close to Harsanyi's extended-preference model. This should make the objectivist model more palatable to the reader accustomed to thinking of welfare in economic terms, as exclusively a function of preference, since Harsanyi's and similar approaches have been the most influential accounts of interpersonal comparison within welfare economics.¹⁵⁶ The crucial difference between Harsanyi's model and my own is this: While Harsanyi simply assumes that idealized observers I_1, I_2, \dots, I_n will have the same preferences as between the S_2 -to- S_1 moves for P and for Q , I make no such assumption. If two observers under idealized conditions of full information, intensive deliberation, and so on, would have a different preference as between the two moves, then the comparative effect of S_1 and S_2 on the overall well-being of P and Q is indeterminate. If and only if all idealized observers would have *conver-*

154. See *id.* at 187 ("My handing back a wallet I found in the street in such and such circumstances is right, for example, only if, under conditions of full rationality, we would all want that if we find a wallet in the street in such and such circumstances, then we hand it back.").

155. This approach, like Harsanyi's approach, would be extended to comparisons involving more than two persons through the device of a lottery. Each idealized observer would ask how she would compare the S_1 lottery (the chance of being each person in S_1 , with probability $1/n$ where there are n persons in the population) to the S_2 lottery (the chance of being each person in S_2 with probability $1/n$). See Weymark, *supra* note 140, at 255 (explaining Harsanyi's approach, including his use of an equiprobability lottery over person-states).

156. In particular, it is the approach adopted by Kenneth Arrow. See, e.g., Kenneth J. Arrow, *Extended Sympathy and the Possibility of Social Choice*, 7 *PHILOSOPHIA* 223 (1978); Kenneth J. Arrow, *Welfare Economics: Extended Sympathy and the Possibility of Social Choice*, 67 *AM. ECON. REV.* 219, 224 (1977).

gent preferences over the moves can a determinate statement about overall welfare be made.

But this is no odder than saying that some person's performance of an action (*A*) is morally better than the alternative only if all idealized observers (or all participants in a hypothetical Rawlsian or Scanlonian scenario of social contracting) would agree that a person in those circumstances should perform *A*. Some of our idealized preferences or judgments *do* converge;¹⁵⁷ and it is just in the case of such convergence that we can speak of objective rightness or goodness, rather than mere subjective preference. The compelling fact patterns that undermine the putative impossibility of interpersonal comparisons—for example, the case where *S*₁ involves a painful death for thousands, while in *S*₂ the thousands do not die, but *P* suffers a mild headache—are just cases in which the comparative effects of the outcomes on persons are sufficiently stark that we expect any thoughtful and well-informed observer to reach the same welfare evaluation of the outcomes.

What criticisms might be leveled against the objectivist approach to interpersonal comparison sketched here? One criticism is that this approach leaves some outcomes welfare-incomparable. Welfare economists often demand that a social welfare function give a complete ranking of states, such that every outcome is better, worse, or exactly equal to every other. The approach sketched here does not do that.¹⁵⁸ If *S*₁ is Pareto-incomparable with *S*₂, and if idealized observers would not converge in their extended preferences over the two outcomes, then, on the objectivist approach, the outcomes are neither better, nor worse, nor equally good with respect to overall welfare.

However, completeness is too strong a feature to require of a moral criterion. It is a welcome feature, but not a necessary one. (As I have already observed, the Pareto criterion is itself incomplete.) Various prominent philosophers and economists—including Joseph Raz,¹⁵⁹ Amartya Sen,¹⁶⁰ Thomas Hurka,¹⁶¹ and Larry Temkin¹⁶²—have persuasively argued that we should accept the existence of some incomparabilities with respect to the criterion of overall well-being or with respect to related criteria (for example, the criterion of equal well-being).¹⁶³ For example, if *P* goes for a revitalizing walk in the

157. See SMITH, *supra* note 92, at 187-89.

158. See, e.g., BOADWAY & BRUCE, *supra* note 29, at 137-70.

159. See JOSEPH RAZ, *THE MORALITY OF FREEDOM* 322-35 (1986).

160. See AMARTYA SEN, *INEQUALITY REEXAMINED* 46-48 (1992).

161. See THOMAS HURKA, *PERFECTIONISM* 86-88 (1993).

162. See TEMKIN, *supra* note 61, at 141-47.

163. For philosophical and legal discussion about the issue of "incommensurability," including (although not necessarily limited to) the problem of incomparability, see *INCOMMENSURABILITY, INCOMPARABILITY AND PRACTICAL REASON* (Ruth Chang ed., 1997). See also ANDERSON, *supra* note 62, at 44-64; GRIFFIN, *supra* note 73, at 75-92; HURLEY,

park in S_1 but not in S_2 , and Q enjoys an evening with friends in S_2 but not in S_1 , and the two outcomes are otherwise identical, it seems very plausible that (1) S_1 is neither better nor worse for overall well-being than S_2 (revitalizing walks and friendly talks do not differ in their effect on the overall welfare calculus); but also (2) S_1 is not precisely equal to S_2 in light of overall well-being (since the outcome S_3 in which P goes for a slightly more revitalizing walk is better for overall well-being than S_1 , but still neither better nor worse than S_2); and therefore (3) S_1 and S_2 are *incomparable* with respect to overall well-being.

A second and more serious objection is that sophisticated preferentialism is inconsistent with an objectivist approach to interpersonal comparison—that this approach would lead us to a more objectivist theory of well-being, either simple objectivism or the more plausible view I have term sophisticated objectivism. But where is the inconsistency? Sophisticated preferentialism is a much better theory than simple objectivism because it preserves the truism that P cannot be benefited by an outcome unless he prefers it (or comes to prefer it). It is perfectly coherent to endorse that truism and, at the same time, to claim that judgments of overall well-being with respect to Pareto-noncomparable states are a function of objective goods.

Think of the point this way: In deciding who is harmed by a social choice (a Loser), benefited by that choice (a Winner), or unaffected by the choice (a Neutral), sophisticated preferentialism relies solely on persons' preferences and idealized preferences; it is only at the *latter* stage of comparing welfare gains to the Winners with welfare losses to the Losers that objective goods come into play. By contrast, simple objectivism ignores preference at every stage, which makes it an implausible account of welfare, while sophisticated objectivism brings actual preference into play at the first stage, but also brings in objective goods (rather than idealized preference) at that initial stage. Whether sophisticated preferentialism or sophisticated objectivism is the correct account of welfare is a close question, but for reasons I elaborated in Part II, I think the basic categorization of persons as Winners, Losers, and Neutrals *should* be solely a function of preference and idealized preference rather than objective goods.

Finally, the objectivist approach to interpersonal comparison that I have advocated may be criticized because it seems illiberal. Is not

supra note 67, at 254-70; JOHN KEKES, THE MORALITY OF PLURALISM 53-75 (1993); ISAAC LEVI, HARD CHOICES (1986); NUSSBAUM, *supra* note 62, at 107; RAZ, *supra* note 159, at 321-22; HENRY S. RICHARDSON, PRACTICAL REASONING ABOUT FINAL ENDS 89-118 (1997); STOCKER, *supra* note 64, at 211-40; DAVID WIGGINS, NEEDS, VALUES, TRUTH 239-62 (1987); Matthew Adler, Introduction to Symposium, *Law and Incommensurability*, 146 U. PA. L. REV. 1169 (1998); John Finnis, *Natural Law and Legal Reasoning*, 38 CLEV. ST. L. REV. 1 (1990); T.K. Seung & Daniel Bonevac, *Plural Values and Indeterminate Rankings*, 102 ETHICS 799 (1992).

the invocation of objective welfare-judgments inconsistent with the liberal claim that government should be neutral between differing ways of life?¹⁶⁴ But sophisticated preferentialism is significantly responsive to the liberal claim, insofar as it is *P*'s own preferences that determine whether he is benefited, harmed, or unaffected by an outcome—whether he is a Winner, Loser, or Neutral. And the definition of objective goods in terms of convergent ideal judgments should further allay the liberal's concern; if well-informed persons genuinely disagree about the worth of the different ways of life at stake, then ideal judgments will not converge. The liberal who demands yet more, who eschews any talk whatsoever of objective welfare-value, will be forced to either: (1) claim that interpersonal comparisons are impossible; (2) claim that the correct approach to interpersonal comparisons is some approach that doesn't involve objective values, such as the ordinal, cardinal, extended-preference, or cost-benefit approaches; or (3) claim that overall well-being lacks moral force. I have already demonstrated why the first two claims are spurious. Let me now try to show why the third one also is spurious.

B. *The Moral Force of Overall Well-Being*

Imagine that a regulatory action, policy, or institution will increase overall well-being relative to alternative regulatory options, and that this welfare-increasing option does not run afoul of deontological requirements, perfectionist values, or norms of distributive justice sufficiently weighty to override the criterion of overall well-being. Then, the welfarist claims, government is morally obliged to choose the welfare increasing option. Welfarism says that overall well-being is morally relevant, if not morally conclusive; the fact that a regulatory option is welfare-increasing constitutes a *prima facie* moral reason for regulators to choose the option. That is, it constitutes a moral reason that *will* eventuate in an all-things-considered obligation to choose the option, *if* there are no overriding moral reasons to the contrary. But why? Why is overall well-being even morally relevant, let alone morally conclusive? Why does that feature of options and outcomes have moral force?

This important question has been generally ignored, both by welfare economists and by moral philosophers.¹⁶⁵ Welfare economists have focused exclusively upon the question addressed in Section A above, namely, is overall well-being determinate?¹⁶⁶ Their working

164. See SHER, *supra* note 74 (discussing and criticizing various liberal theories, e.g., those of Ackerman and Dworkin).

165. Shelly Kagan is one exception. See KAGAN, *supra* note 98, at 29-40 (arguing that well-being has *prima facie*, if not conclusive, moral relevance).

166. See, e.g., Robbins, *supra* note 55; Pollak, *supra* note 56; INTERPERSONAL COMPARISONS OF WELL-BEING, *supra* note 134.

and undefended assumption seems to be that, if overall well-being is determinate—if interpersonal comparisons are indeed possible—then overall well-being (“social welfare”) or something like it ought to play a role in government decisionmaking.¹⁶⁷ The assumption is problematic. It is a necessary condition for overall welfare to possess moral force that it be determinate, but it is not a sufficient condition; I can concede that my losses can be compared with your gains, without conceding that your gains constitute a moral reason (even a *prima facie* one) for me to suffer losses. This is a point well recognized within the philosophical literature. Philosophers have generally assumed the interpersonal comparability of welfare and have instead trained their attention on the question, Is it morally obligatory to maximize overall well-being?

But that question is almost always mooted within the context of a debate about utilitarianism.¹⁶⁸ Utilitarians claim that it is always morally obligatory, all things considered, to maximize overall well-being. In other words, they argue that overall well-being is morally *conclusive* and not merely morally relevant. The critics of utilitarianism advance familiar counterexamples (for example, killing one to save five, sacrificing the scapegoat to appease the crowd) to show why welfare maximization is not always morally permissible and, *a fortiori*, why it is not always morally obligatory. Although these counterexamples are persuasive, they leave untouched the welfarist’s weaker claim that government officials are subject to a *prima facie* moral requirement to engage in welfare maximization. Yet the claim is hardly self-evident; *pace* the welfare economists, we need an argument for the moral relevance of “social welfare.”

What would that argument be? The welfarist faces two potent objections. Call the first the “minimalist” objection; call the second the “egalitarian” objection. The minimalist objection is that no person has a claim that others suffer losses so as to improve her welfare (at least if those others lack a special relationship to her).¹⁶⁹ The egali-

167. For a good illustration of this assumption, see Richard Layard and Stephen Glaister, *Introduction to COST-BENEFIT ANALYSIS*, *supra* note 29, at 1-2. They reason:

If we have to decide whether to do A or not, the rule is: Do A if the benefits exceed those of the next best alternative course of action, and not otherwise. . . . [N]o one could complain at that. . . . But how are [numerical] values [of costs and benefits] to be arrived at? If we assume that only people matter, the analysis naturally involves two steps. First, we must find out how the decision would affect the welfare of each individual concerned. . . . The second step is to deduce the change in social welfare implied by all the changes in individual welfare.

Id.

168. For some entries in, or overviews of, this debate, see SCARRE, *supra* note 128; J.J.C. SMART & BERNARD WILLIAMS, *UTILITARIANISM: FOR AND AGAINST* (1973); *UTILITARIANISM AND BEYOND 1* (Amartya Sen & Bernard Williams eds., 1982).

169. See SHELLY KAGAN, *THE LIMITS OF MORALITY* 5 (1989) (delineating “minimalist” view); ROBERT NOZICK, *ANARCHY, STATE AND UTOPIA* (1974) (defending libertarianism, the standard variant of minimalism).

tarian objection is that no person has a claim that others suffer losses so as to improve her welfare just by virtue of the fact that her welfare gain is larger than their losses; rather, her justified claim is that others suffer losses so as to provide her a fair (equal) amount of welfare.¹⁷⁰

Welfarism and Its Critics

Welfarism: Where a regulatory action produces a gain for person P_w and a loss for P_l , such that P_w 's gain is larger than P_l 's loss, then government has a moral reason to perform the action and will be morally required to perform it unless other (and overriding) moral considerations are at stake.

Minimalism: Where a regulatory action produces a gain for person P_w and a loss for person P_l , such that P_w 's gain is larger than P_l 's loss, then government has no moral reason yet to perform the action. (In other words, government will not be morally required to perform the action, even if moral considerations besides overall well-being are not at stake.) Similarly, where a regulatory action produces a gain for P_w and a loss for P_l , such that this action lessens the inequality between P_w and P_l (because P_w is poorer than P_l), then government has no moral reason yet to perform the action. (Government will not be morally required to perform it even if moral considerations besides equality are not at stake.) P_w 's welfare is simply not a sufficient basis for imposing a welfare setback upon P_l . Although P_l is subject to moral duties that are, in some sense, grounded in P_w 's welfare (for example, the duty not to intentionally and directly kill P_w , to breach a promise to him, or to deprive him of his property), the sheer fact that P_w would be better off is no grounds whatsoever for making P_l worse off.

Egalitarianism: Where a regulatory action produces a gain for person P_w and a loss for P_l , such that P_w 's gain is larger than P_l 's loss, whether government has a moral reason to perform this action is essentially dependent on the welfare levels of P_w and P_l and of other persons in the population. If the action lessens the inequality of welfare levels (for example, if everyone else is at level M , and the action moves both poorer P_w and richer P_l closer to level M), then government has a moral reason to perform it. If the action increases the inequality of welfare levels (for example, if everyone else is at level M , and P_w is already above it while P_l is below it), then government has a moral reason not to perform the action.

170. In recent years, numerous moral philosophers have advanced some kind of egalitarian view. See, e.g., ERIC RAKOWSKI, EQUAL JUSTICE (1991); TEMKIN, *supra* note 61; Arneson, *supra* note 67; G.A. Cohen, *Where the Action Is: On the Site of Distributive Justice*, 26 PHIL. & PUB. AFF. 3 (1997); Ronald Dworkin, *What Is Equality? Part 2: Equality of Resources*, 10 PHIL. & PUB. AFF. 283 (1981). See generally JOHN E. ROEMER, THEORIES OF DISTRIBUTIVE JUSTICE (1996). Clearly, not all of these theorists would defend the specific egalitarian view spelled out below; but that view fits comfortably within the family of egalitarian theories just cited.

Further, governmental actions will generally increase or decrease the equality of welfare, relative to alternative actions. Although overall well-being might be a relevant consideration when choosing between options that are equally or incomparably good with respect to welfare-equality, that is a highly unusual case.

The welfarist needs to take very seriously both the minimalist objection and the egalitarian objection. In effect, these objections represent two horns of a dilemma that welfarism must navigate. Either we are sufficiently linked by the sheer fact of coexistence in a common society (or on a common planet) that everyone has a claim on everyone else to a fair and equal share of social welfare (which everyone can rightly claim to have helped produce);¹⁷¹ or we are not thus linked,¹⁷² in which case fellow citizens or Earthlings are under no mutual moral obligation to equalize welfare, maximize welfare, or otherwise take account of welfare effects on each other.

I have no foundational rebuttal to the minimalist and egalitarian objections. A foundational argument for welfarism would try to show why, given certain plausible premises about the foundations of morality, overall well-being turns out to have moral force. For example, the welfarist could argue (1) that fellow citizens or Earthlings are linked in a common scheme of cooperation—*pace* the minimalist—such that the welfare each reaps is a joint product of everyone's effort; (2) that in such a situation, moral obligations are those obligations that the citizens or Earthlings would choose in a initial contract setting out the terms of social cooperation;¹⁷³ and (3) under the proper specification of the contracting scenario (for example, as a scenario where contractors do not know what the particulars of their lives will be), the citizens or Earthlings would agree to welfarism, *pace* the egalitarian. But I am not sure that this particular foundational argument for welfarism works.

Harsanyi has famously tried to demonstrate that social contractors bargaining behind a veil of ignorance would indeed agree to maximize well-being,¹⁷⁴ rather than (as Rawls claims) to maximize the position of the least well-off, or to choose an egalitarian standard. Yet Harsanyi's demonstration is controversial,¹⁷⁵ both because it relies upon controversial assumptions about the nature of rational choice under uncertainty and, even more importantly, because it constitutes an argument for utilitarianism, which as I have explained is vulnerable to persuasive counterexamples. What we would need is a

171. See RAWLS, *supra* note 99.

172. See NOZICK, *supra* note 169.

173. See KAGAN, *supra* note 98, at 240-56 (describing "contractarian" theories about the foundations of morality).

174. See Weymark, *supra* note 140, 291 (presenting Harsanyi's argument).

175. See *id.*

hypothetical-contract argument for welfarism, but not for utilitarianism; it is unclear to me how that argument should be developed. More generally, it is unclear to me that a deductive argument from any plausible view about moral foundations will lead us to welfarism but not to utilitarianism. Welfarism is a *pluralist*, not a monist view; it gives force to one moral item, while leaving space for qualitatively distinct others (rights, equality, perfections). Pluralist moral views are notoriously resistant to foundational demonstration.

Rather, my argument for welfarism will appeal to moral intuitions.¹⁷⁶ It is counterintuitive to think that overall well-being lacks *prima facie* moral force, or so I will claim. The counterintuitive implications of minimalism and egalitarianism emerge quite clearly in the following kind of case.

The Beneficial Meteor: A Counterexample to Minimalism and Egalitarianism

A shower of meteors has landed on government property. Most of the rocks are needed for scientific research and museums, but there is one extra one. This meteor is useless, except for the following: (1) it can be used to relieve P_w 's emphysema, a chronic (but non-fatal) condition that causes him near-constant discomfort and prevents him from engaging in sports or other vigorous activities; and (2) it can be used to alleviate a very mild headache that P_l suffers once every month, which like P_w 's emphysema is resistant to other medical treatments. P_w is much wealthier than P_l . P_w is a millionaire. P_l is not impoverished, but neither is she wealthy or even close. Her income and wealth place her in the lower middle class, at around the twenty-fifth percentile, while P_w 's income and wealth place him well above the ninety-fifth percentile.

What should government do with the beneficial meteor? I suggest that government is morally required to use the meteor to alleviate P_w 's emphysema—in short, to choose the option that maximizes welfare. First, it seems intuitively clear that government is morally required either to use the meteor to help P_w or to use it to help P_l . Destroying the meteor is a third option for government, but that would be wrong; the rock can make a significant difference to P_w 's life and a smaller, but still observable, difference to P_l 's. Further, as between the options of giving the meteor to P_w and giving it to P_l , it is the welfare-maximizing option and not the welfare-equalizing option that seems morally correct in this case. Government has a *prima facie* moral reason to give the meteor to P_w because it matters to his welfare; government also has a *prima facie* (and conflicting) moral reason to give the meteor to P_l because it matters to *her* welfare. Yet the

176. See DANIELS, *supra* note 149 (describing the “reflective equilibrium” approach to moral reasoning, which relies in part upon moral intuitions).

effect on P_I 's welfare of giving the rock to her is comparatively quite trivial, as compared to the effect upon P_w 's welfare of giving it to him. Relatedly, the (negative) effect upon welfare equalization caused by giving P_w the meteor is seemingly small, as compared to its (positive) effect upon overall welfare. Thus, I propose, it is P_w and not P_I who ought to get the meteor.

The reader may not share my intuitions about the Beneficial Meteor case. She may have strong minimalist leanings, and conclude that it is indeed a matter of moral indifference whether government destroys the rock, as opposed to giving it to P_w or P_I . Or, she may have strong egalitarian leanings, and conclude that government is morally prohibited from either destroying the rock or giving it to P_w . Instead, she may insist, the gains to equality that flow from alleviating P_I 's mild headache *do* take priority over the gains to aggregate welfare that flow from alleviating P_w 's emphysema. I have no systematic or foundational argument to show that such views about the Beneficial Meteor case are incorrect. If the reader has these views, then she can quite happily deny that welfarism is the right moral theory. Appeal-to-intuition arguments within moral philosophy are persuasive only insofar as the posited intuitions are indeed shared. "But that's *not* my reaction" is always a possible response when the philosopher claims that a particular reaction to a given case is natural or obvious. On the other hand, if the reader does share my reaction to the Beneficial Meteor case, that the rock should go to P_w , then she should also espouse welfarism. That reaction and a denial of welfarism are inconsistent.

Is this correct? Is it really true that the joint posture of (1) having government choose the welfare-maximizing option in Beneficial Meteor and other cases like it and (2) denying welfarism is untenable? One way to defend the choice of the welfare-maximizing option in Beneficial Meteor, while denying welfarism, is to argue that this option, giving the meteor to P_w rather than P_I , is actually justified on some basis other than its maximization of welfare. But what would that basis be? The moral requirement (such as it may be) that welfare or its prerequisites be equally distributed cannot be the basis for the choice of P_w . P_w is much wealthier than P_I , and his welfare level is much higher, notwithstanding the misfortune of emphysema. Nor do considerations of moral desert¹⁷⁷ ground that choice. P_w might have performed worthy actions in the past, for which he now deserves a reward (or reciprocally P_I might have done past wrong, for which her nonreceipt of the meteor is fair punishment, trumping her egalitarian claim to it); but the facts of Beneficial Meteor, as I have stated them, do not entail that P_w and P_I are differentially situated

177. See generally GEORGE SHER, DESERT (1987).

with respect to considerations of moral desert. The intuition that needs to be explained is why P_w should get the rock apart from such considerations.

Perhaps the basis for P_w 's claim is that he has a greater *physical need* for the meteor than P_i ? The notion here, as articulated in the work of Thomas Nagel and Tim Scanlon, is that well-being *per se* is morally inconsequential. What stand in its stead, so Scanlon and Nagel claim, are certain elements of well-being, such as physical pains and pleasures:

If you and a stranger have both been injured, you have one dose of painkiller, and his pain is much more severe than yours, you should give him the painkiller—not for any complicated reason, but simply because of the relative severity of the two pains, which provides a neutral reason to prefer the relief of the more severe. . . .

But many values are not like this. Though some human interests (and not only pleasure and pain) give rise to impersonal values, I now want to argue that not all of them do. If I have a bad headache, anyone has a reason to want it to stop. But if I badly want to climb to the top of Mount Kilimanjaro, not everyone has a reason to want me to succeed. I have a reason to try to get to the top, and it may be much stronger than my reason for wanting a headache to go away, but other people have very little reason, if any, to care whether I climb the mountain or not.¹⁷⁸

The Scanlon-Nagel view might be fleshed out as follows: (1) Only certain elements of well-being (call them “needs”) have moral force. If I lack a special relationship to you, I have moral reason to satisfy your needs, but not to increase your well-being *per se*. (2) Where needs conflict, we should maximize the satisfaction of needs, *ceteris paribus*. If one option is satisfying some of P_1 's needs by amount N_1 , and another option is satisfying some of P_2 's needs by amount N_2 , and N_1 is larger than N_2 , then we should choose the first option, *ceteris paribus*. (3) In short, we have moral reason to maximize need satisfaction, but not to maximize well-being. (4) Giving the meteor to P_w in the Beneficial Meteor case is justified because it maximizes need satisfaction—because P_w 's need for relief from emphysema is greater than P_i 's need for relief from a mild headache—and not because it maximizes well-being.

The difficulty with this account of the Beneficial Meteor case is that the moral force of needs and the moral force of welfare turn out on closer examination to be inseparable. A harmless need, that is, one that involves no welfare-setback for the person in need, surely

178. Thomas Nagel, *Autonomy and Deontology*, in CONSEQUENTIALISM AND ITS CRITICS 142, 145-46 (Samuel Scheffler ed., 1988). For similar claims by Scanlon, see his book cited *supra* note 101.

lacks moral force. Imagine that P_w is a stoic who sees the emphysema as a test of his fortitude and thus prefers to have the condition and would retain this preference under ideal conditions. Then there is no reason to give him the meteor. Scanlon and Nagel might concede the point—they might even concede that needs are necessarily harmful, that conditions of a person do not really constitute “needs” unless they reduce his welfare—yet still insist that only needs, rather than harmful conditions of other kinds, give rise to moral requirements. But why should this be the case?

Imagine that P_w is replaced by P_w^* , who is not suffering from emphysema and therefore does not *need* the meteor, but who *can* use it to prevent some welfare-setback that (in terms of aggregate welfare) is still larger than P_i 's headache. For example, suppose P_w^* has labored for years to develop a unique and beautiful garden and all the plants have contracted a rare disease for which the meteor's minerals provide the only cure. If P_w 's emphysema matters morally only because it harms him, and the loss of P_w^* 's garden also harms him, and both harms are larger than the trivial harm to P_i comprised by his headache, then why should P_w , but not P_w^* , get the meteor?¹⁷⁹ Perhaps it is generally true that relief from physical suffering (emphysema, headaches) is more important for welfare than nonneeded benefits (the flowering of a garden). However, in a case like the just-described modification to Beneficial Meteor where the generalization is untrue, it is the bedrock moral criterion of overall welfare, not the proxy of need satisfaction, that determines what choice is morally required.

Another plausible way to argue that the choice of the P_w option in the Beneficial Meteor case is actually justified on some basis other than welfare maximization is as follows:

The Prioritarian Account of Beneficial Meteor

Government is required, not to maximize overall well-being, that is, the unweighted sum of individual well-being, but rather to maximize the sum of each individual's well-being as weighted by a factor that gives greater weight to the well-being of those worse off, and thus is inversely proportional to that individual's level of wealth or welfare. Let us say that P_i^O is the welfare of each person (P_i) in outcome O . Then, in comparing outcomes O and O^* for the set of persons $\{P_1, P_2, \dots, P_n\}$, what is morally relevant is not $[(P_1^O - P_1^{O^*}) + (P_2^O - P_2^{O^*}) + \dots + (P_n^O - P_n^{O^*})]$, but rather something like

179. James Griffin provides a similar illustration:

A group of scholars may, with full understanding, prefer an extension of their library to exercise equipment for their health. . . . But then to maintain that needs create obligations where mere desires do not, or that they create stronger obligations, is to say that we have an obligation, or a stronger one, to the scholars to give them what they themselves value less, which would be odd.

GRIFFIN, *supra* note 73, at 45.

$[d(P_1)(P_1^O - P_1^{O*}) + d(P_2)(P_2^O - P_2^{O*}) + \dots + d(P_n)(P_n^O - P_n^{O*})]$, where $d(P_i)$ is inversely proportional to the level of each P_i 's welfare or wealth. Giving the meteor to the emphysema sufferer, P_w , in the Beneficial Meteor case is justified only because the effect of the meteor on his welfare ($P_w^O - P_w^{O*}$) is sufficiently greater than the effect of the meteor on P_i 's welfare so as to outweigh the fact that $d(P_w)$ is smaller than $d(P_i)$.

A prioritarian view of morality, one that compares outcomes with reference to the criterion of weighted overall well-being, rather than overall well-being *simpliciter*, has been advanced in the philosophical literature¹⁸⁰ and, to some extent, in the modern welfare-economics literature.¹⁸¹

The problem is that any set of weights posited by the prioritarian will end up seeming arbitrary. Given any set of weights, a large enough improvement in the welfare of a wealthier person (P_w) will justify, at least *prima facie*, a particular welfare setback to a poorer person (P_i). Why, then, should not a smaller improvement that increases aggregate welfare also justify, at least *prima facie*, that same welfare setback to the poorer person? The prioritarian might answer that the criterion of overall well-being must be balanced against the claims of equality or fair distribution—that the weighted aggregate criterion is what emerges when we start with the more basic criteria of (1) unweighted aggregate welfare and (2) equal welfare, or some such basic distributive criterion. I am skeptical that a weighted criterion would thus emerge from a balancing of unweighted aggregate welfare and a basic distributive criterion. A balancing of those criteria will, I think, lead to some incomparabilities that the prioritarian's weighted criterion will fail to reflect.¹⁸² In any event the defense of prioritarianism just mooted effectively concedes that overall well-being *does* have *prima facie* moral force.

Alternatively, the prioritarian might answer that the weights just emerge from the right specification of morality's foundations, for example, from the right specification of the social contracting scenario. However, the best developed analysis along these lines, Rawls' argument in a *Theory of Justice*, leads not to the weighted aggregate

180. See SAMUEL SCHEFFLER, *THE REJECTION OF CONSEQUENTIALISM* (rev. ed. 1994). Shelly Kagan has suggested:

Perhaps, then, to the extent that distribution matters, this factor should be captured in terms of a principle that gives greater weight to improving the well-being of those who are worse off: the lower the level of someone's well-being (in absolute terms), the greater the extent to which increasing their well-being by a certain amount improves the goodness of an outcome.

KAGAN, *supra* note 98, at 53.

181. See, e.g., BOADWAY & BRUCE, *supra* note 29, at 137-70 (presenting various types of social welfare functions, including those that incorporate a weighted rather than unweighted sum of individual utilities).

182. See *supra* notes 158-59 and accompanying text (discussing incomparability).

criterion but rather to a “maximin” criterion that gives lexical priority to the well-being of the worse off.¹⁸³ If the prioritarian wants to justify the choice of P_w in the Beneficial Meteor case, and I am assuming here that she does, the maximin criterion will not provide that justification.

Finally, the prioritarian could claim that it is weighted, not unweighted, aggregate welfare which corresponds to (one component of) the overall “value” or “goodness” of outcomes. A suitably detached observer, evaluating the “goodness” of world-states, would take weighted rather than unweighted aggregate welfare to be one dimension of goodness.

However, this claim seems implausible. The prioritarian agrees that the relevant dimension of goodness is tightly linked to human welfare; it is human flourishing, or some function thereof, that is at stake. The prioritarian also agrees that human flourishing or welfare is interpersonally comparable. (Indeed, his weighting scheme assumes that each person’s welfare in a given outcome can be represented by an interpersonally valid number, a “utility,” which is weighted and then summed with the weighted utilities of other persons.) The prioritarian should further agree that, as between an outcome where overall well-being is greater and an outcome where weighted aggregate welfare is greater, there is more welfare or flourishing in the first outcome. (Consider this analogy: The world with more mass is the world where the aggregate mass of the objects in it is greater.) Finally, the prioritarian himself is concerned with an aggregative kind of goodness—the kind of goodness that is connected to “more” or “less.” (Note that a change that produces dramatically more welfare for one person and leaves everyone else unaffected will be counted by the prioritarian as a *better* world, regardless of the fact that the distribution of welfare may be less equal, balanced, or harmonious.) All this forces the prioritarian to claim, implausibly, that the world with more welfare or flourishing is not—even from a suitably detached perspective, and even when an aggregative kind of goodness is at stake—the better world, as far as welfare is concerned.

So much for the strategy of choosing P_w in Beneficial Meteor while denying welfarism by proposing some criterion other than overall well-being that purportedly justifies choosing P_w . A second and quite different way to defend the choice of the welfare-maximizing option P_w in Beneficial Meteor, while disclaiming a welfarist theory of regulation, is to argue that government’s choices in Beneficial Meteor are disanalogous to the choices that government regulators face. The task of regulators is not to parcel out the meteors that fall on the public lands, or to distribute other sorts of public property, but pri-

183. RAWLS, *supra* note 99.

marily to impose legal duties which constrain the actions of private citizens. Individuals are commanded not to pollute, not to endanger species, not to collude in setting prices, not to impose certain risks upon their employees, and so on. Assume that *D* is some duty-creating (act-constraining) directive, and that a regulator is deciding whether to issue *D*. If government ought to choose the option that maximizes overall well-being in Beneficial Meteor, but ought *not* choose between *D* and *not-D* on the basis of overall well-being, then the lessons of Beneficial Meteor for regulators are indeed very limited.

Why would one think that overall well-being has different force with respect to duty imposition than it does with respect to meteor-distribution? The plausible idea here is that duty imposition raises deontological issues that are not present in Beneficial Meteor. To be concrete, imagine that *P_i* is the only person practically constrained by *D*. Also imagine that the issuance of *D* will cause *P_i* a slight welfare harm (equivalent to the harm of a mild headache), and that the failure to issue *D* will produce a large welfare setback for *P_w* (equivalent to the harm of emphysema). If the issuance of *D* will infringe some deontological constraint protecting *P_i*—if it will infringe her “rights”—then the sheer fact that *D* maximizes welfare will not justify that option. Welfare maximization, without more, does not warrant the breach of deontological constraints.¹⁸⁴ Further, the view that any legal duty, at least any legal duty backed by sanctions, implicates deontological constraints (because such duties are coercive restrictions of liberty) is a possible one.

I do not have space here to show why this robustly deontological view of regulatory duties is wrong. Let me simply point out that the view is radically at odds with the practices of regulatory agencies. Roughly, if *D* infringes the right to liberty, then *D* is justified only if the actions it constrains are themselves rights-violating. However, many of the actions constrained by the antitrust laws, the environmental laws, the food and drug laws, the workplace safety laws, and other regulatory regimes are not rights-violating actions. Rather, they are simply actions that produce bad consequences, by reducing well-being, increasing inequality, degrading perfectionist values (such as environmental values), or producing some other undesirable result. There is generally no deontological violation in directing persons to refrain from such actions—surely not if the directive is civil rather than criminal.¹⁸⁵ Assume this is correct; again, if it is not, then the proper scope of regulation is radically more limited than most

184. See sources cited *infra* note 189.

185. Cf. JOEL FEINBERG, HARM TO OTHERS 26 (1984) (articulating the view that a behavior should not be criminalized unless it is harmful or seriously offensive to some person other than the actor).

modern theorists (neoclassicists, proceduralists, and others) have thought. Therefore, the imposition of D on P_l does not implicate deontological constraints. Then P_w 's claim to the issuance of D is just as strong as his claim to the beneficial meteor. Both D and the meteor help P_w much more than they harm P_l . Both cause only a slight setback to welfare-equality, and both are deontologically permissible. Overall well-being gives the government sufficient moral grounds to issue D , just as it gives government sufficient grounds to give P_w the meteor. The Beneficial Meteor case has plenty of relevance for regulation.

C. Deontology, Perfections, and Distribution: A Note on Nonwelfarist Considerations

As I have repeatedly emphasized, welfarism refers to the family of moral theories that make overall well-being morally relevant, not necessarily morally conclusive. Welfarism, thus defined, includes but is not limited to utilitarianism. Further, as I have suggested, there are strong arguments against utilitarianism. The *right* welfarist theory, the theory that *correctly* evaluates the actions of regulators and other government officials, will be a theory such that the moral status of a regulatory action is a function of a plurality of moral criteria (W^* , X_1 , X_2 . . . X_n , where W^* is overall well-being and X_i is another moral criteria). Unless W^* takes lexical priority over the X_i , which is little more plausible than utilitarianism, welfare-maximizing regulatory actions may turn out to be morally wrong. The right variant of welfarism will acknowledge that "overall well-being" can be outweighed by other moral considerations.¹⁸⁶

What, more specifically, are the conditions under which welfare-maximizing regulatory actions will turn out to be morally wrong? This Article does not purport to address that crucially important issue. It is plausible that a welfare-maximizing regulatory action can be morally wrong, all things considered, because it infringes deontological constraints, reduces overall perfection, or reduces overall equality or the overall satisfaction of other such "distributive" criteria. But which deontological constraints, perfectionist goals, and distributive goals really do bear upon the moral status of regulation is a large and unanswered question, which I do not have the space to consider in any detail here. Nor do I mean to insist that the only criteria relevant to government regulation, other than the criterion of overall well-being, are deontological, distributive, or perfectionist criteria. I simply cite these as the most plausible possibilities. A theory incorporating W^* plus some esoteric X_i that is neither deontological,

186. On pluralism, see, for example, KEKES, *supra* note 163; STOCKER, *supra* note 64; SUMNER, *supra* note 9, at 200 n.24 (citing sources).

distributive, nor perfectionist would still be a welfarist theory as I have defined “welfarism.”¹⁸⁷

Deontological constraints are nonconsequentialist duties—generally duties of inaction rather than action.¹⁸⁸ They would prohibit regulators from performing certain types of actions, even where the consequences of such actions are better than the consequences of inaction. The pros and cons of deontological constraints have been much discussed within the modern philosophical literature.¹⁸⁹ I tend to think that the deontologists have made a strong case for the existence of some such constraints—for example, a constraint against direct, intentional killing (I ought not directly and intentionally kill one person, even to prevent five direct and intentional killings). Assuming this is true, it remains possible that deontological criteria and consequentialist criteria are “partitioned” between different governmental institutions—between courts, legislatures, regulators, executives, state institutions, and federal institutions—such that regulators are effectively permitted to ignore deontological criteria.¹⁹⁰ In particular, the following idea is appealing:

The Constitutional (Deontological) Partition:

Every deontological constraint binding upon regulators is incorporated in some justiciable requirement of constitutional law. For example, it would violate the substantive component of the Due Process Clause for a regulatory official to perform a direct and intentional killing.¹⁹¹ As between two or more *constitutional* options, the morally appropriate option for the regulator to choose is the option with the best consequences.

The Constitutional Partition is appealing because it restores simplicity to the regulator’s moral world: once constitutional courts are in place, the regulator can act as if morally right actions are, simply, actions with the best consequences.

But is the Constitutional Partition true? If regulators were deontologically constrained, not merely to refrain from consequentially

187. But note that I have given “perfectionism” a definition sufficiently broad that it becomes hard to see what that esoteric X_i could be. See *infra* note 194.

188. See David McNaughton & Piers Rawling, *Agent-Relativity and the Doing-Happening Distinction*, 63 PHIL. STUDIES 167 (1991) (providing a precise definition of deontological constraints).

189. See KAGAN, *supra* note 98, at 152-70; NOZICK, *supra* note 169, at 26-53; SCHEFFLER, *supra* note 180, at 80-114; Richard Brook, *Is Smith Obligated that (She) Not Kill the Innocent or that She (Not Kill the Innocent): Expressions and Rationales for Deontological Constraints*, 35 S. J. PHIL. 451 (1997); Heidi M. Hurd, *What in the World Is Wrong?*, 5 J. CONTEMP. LEGAL ISSUES 157 (1994); Frances M. Kamm, *Non-Consequentialism, the Person as an End-In-Itself, and the Significance of Status*, 21 PHIL. & PUB. AFF. 354 (1992); Nagel, *supra* note 178.

190. Cf. Henry S. Richardson, *Institutionally Divided Moral Responsibility*, 16 SOC. PHIL. & POLICY 218 (1999).

191. See *Sacramento v. Lewis*, 523 U.S. 833, 848-50 (1998).

justified killing, but also to refrain from any kind of consequentially justified coercion, the Partition would be clearly false. This is because most if not all regulatory coercion escapes constitutional review, at least under modern (post-*Lochner*) constitutional doctrine.¹⁹² However, as I suggested in Section B above, the intuitive support for a general no-coercion constraint on regulators is much weaker than the intuitive support for a no-killing constraint. Much regulatory coercion—for example, coercive orders directed at corporate officials and backed by civil fines that prohibit pollution, unsafe workplace conditions, or unfair labor practices—is seemingly justified as long as it is consequentially justified.

Yet the failure of the robust deontological view that includes a general no-coercion constraint does not mean that Constitutional Partition is true. The correct deontological account might be a modestly robust view, which does not include a general no-coercion constraint, but which does include the core constraint against direct and intentional killing and also includes other constraints, such as a constraint against the coercive use of the criminal law,¹⁹³ which are not fully incorporated in justiciable constitutional norms. The truly hard, interesting, and unanswered question about deontological constraints for purposes of regulatory theory is whether the correct deontological account is really modestly robust in this way. Constitutional Partition is appealing and, to some extent, plausible, but it still may turn out to be wrong.

Perfectionist criteria and distributive criteria, like the criterion of overall well-being, are species of consequentialist criteria. An action is consequentially justified if the outcome it produces—technically, the total world-state that it produces—is better than the outcomes produced by alternative actions. Consequentialist criteria measure the goodness of outcomes. “Perfectionism”—as I’m using that term here—is the view that features of outcomes other than welfare can matter to the outcome’s goodness.¹⁹⁴ A perfectionist criterion X_i is a criterion such that O_1 can fare better than O_2 with respect to X_i , even though O_1 and O_2 are just the same with respect to welfare, that is,

192. See, e.g., *Ferguson v. Skrupa*, 372 U.S. 726, 731 (1963); *Williamson v. Lee Optical*, 348 U.S. 483, 490-491 (1955).

193. See FEINBERG, *supra* note 185.

194. See generally HURKA, *supra* note 161; SUMNER, *supra* note 9, at 193-95, 208-17. This is a simple, but very broad definition of perfectionism. Arguably, the term is better used to denote a *subset* of the consequentialist moral criteria that depend on features of outcomes other than welfare. For example, one might say that perfectionist criteria, properly speaking, concern the realization of certain capacities essential to humans and other creatures; and one might distinguish between such criteria and *aesthetic* criteria or *environmentalist* criteria, which involve neither welfare nor the realization of essential capacities. Because my project in this Article is to defend welfarism—not to advance a particular view as to the correct moral criteria other than overall well-being—I employ the simple and broad definition of “perfectionism.”

even though the welfare position of every person (and every other welfare subject) in O_1 is just the same as her position in O_2 . Environmental ethicists who argue that the flourishing of nonsentient organisms (organisms without a welfare), endangered species, or whole ecosystems is a relevant consideration for environmental policy, quite apart from human well-being or the well-being of other welfare-subjects, are perfectionists.¹⁹⁵ So are aesthetes, who think that beautiful buildings, objects, paintings, and other artworks contribute to the beauty and thus the goodness of the world even if no one cares about, or otherwise is benefited by, this aesthetic improvement. So is the Aristotelian who thinks that an action which makes a given person more perfect but less happy (for example, by making the person more fully rational, thus creating a more perfect specimen of the species "homo sapiens") is a good action in at least one dimension.¹⁹⁶

The problem with perfectionism, as Wayne Sumner explains, is that:

In doing good for someone we are therefore not merely making the world a better place but also doing something *for that particular person*. Now this notion of furthering someone's good for her own sake can be generalized beyond the boundaries of our species; I have a perfectly good sense of what it means to take my cat to the veterinarian for her sake. But it cannot be generalized indefinitely, for not all natural objects have a "sake" for which we can do things. . . . Our ethical sensibilities seem to have much to do with our ability to see things from the point of view of potential victims and beneficiaries. I am prepared to think that mountains and stars can fare better or worse on some objective scale of perfection, but this fact does not give them a point of view on whose behalf I can marshal my services.¹⁹⁷

I am inclined to find Sumner's arguments against perfectionism persuasive, but the issue is one that merits much further work. The viability of perfectionism has been much less explored by philosophers and theorists than, say, the existence of deontological constraints, or the nature of fair distribution. Note that if perfectionist goals *do* obtain, if they do figure in the best moral theory of government, then it is hard to see how such considerations can be "partitioned" away from regulators.

195. This is the perfectionist view most widely held now, and thus it is the one most relevant to regulatory theory. See, e.g., ROBIN ATTFIELD, *THE ETHICS OF ENVIRONMENTAL CONCERN* (1983); HOLMES ROLSTON, III, *ENVIRONMENTAL ETHICS: DUTIES TO AND VALUES IN THE NATURAL WORLD* (1988); PAUL W. TAYLOR, *RESPECT FOR NATURE* (1986).

196. See HURKA, *supra* note 161.

197. SUMNER, *supra* note 9, at 211.

The Perfectionist Partition:

Every perfectionist goal applicable to regulators is incorporated in some nonregulatory institution, which effectively permits regulators to ignore such goals. For example, if one environmental policy is better in light of overall well-being than a second policy, then regulators should enact the first policy. If it turns out to be all-things-considered worse than the second policy, in light of perfectionist considerations, then it will be reversed by the legislature.

Perfectionist Partition is appealing, because it further simplifies the moral world of regulators. However, I cannot see why it would be true.

Finally, distributive criteria are consequentialist criteria that *do* supervene upon the welfare characteristics of outcomes (if the welfare of every person is just the same in O_1 and O_2 , then O_1 and O_2 must be the same with respect to distributive criteria) but concern the distribution of welfare rather than the aggregate amount.¹⁹⁸ By this definition, the following are all distributive criteria:

- (1) a criterion that measures the extent to which persons are below a certain minimum level of welfare (the "poverty line");
- (2) a criterion that measures the extent to which the overall pattern of welfare levels deviates from perfect equality;
- (3) a maximin criterion, which gives lexical priority to the welfare of the worst-off persons.

All of the above criteria are plausible. John Rawls famously argues in favor of the maximin variant of the idea of proper distribution.¹⁹⁹ Numerous modern philosophers, including Ronald Dworkin, Richard Arneson, G.A. Cohen, and Eric Rakowski, have argued in favor of the egalitarian variant, or something like it.²⁰⁰ Finally, our actual system

198. This definition of "distributive" criteria, like my definition of "perfectionist" criteria, is somewhat simplified. There is a family of views that look to the distribution of "primary goods," of things that serve as *resources* for welfare, rather than to the distribution of welfare itself. See Adler, *supra* note 117, at 1474-76 (discussing these views, and citing sources). A criterion that evaluates outcomes in light of the distribution of welfare-resources is not, strictly speaking, a criterion that supervenes on the welfare characteristics of outcomes, but still it ought to be counted as a "distributive" criterion. (Supervenience fails because, as between two outcomes identical with respect to welfare, one but not the other might have resulted from a fair initial distribution of resources.) It is not clear to me how to modify the simple definition of "distributive" to include resourcist views. The reader should understand that some such modification is required; because this Article is not intended to defend distributive views or to make definitive claims about whether distribution and overall welfare can be "partitioned," I do not pursue the definitional issue here.

199. See RAWLS, *supra* note 99.

200. See sources cited *supra* note 170. It bears noting that some of these egalitarian philosophers are resourcists. See *supra* note 198.

of state redistribution, the “welfare” system, seems to aim at the poverty-line variant.²⁰¹

Here, as with perfectionism and deontology, it is tempting to think that nonregulatory institutions can be structured so that regulators are freed to focus exclusively on aggregate welfare. Some economists working within the neoclassical tradition have claimed that although both efficiency and fair distribution ultimately matter to the evaluation of government actions, it is more efficient and just as good distributively to pursue distributive goals through the tax system.²⁰² Does the claim carry over from efficiency to welfare? Call this the Distributive Partition:

The Distributive Partition:

Every distributive goal applicable to regulators is incorporated in some nonregulatory institution, which effectively permits regulators to ignore such goals.

I find the Distributive Partition no more plausible than the Perfectionist Partition. Consider the version of fair distribution that simply seeks to combat poverty, and is indifferent to differences in welfare levels among persons above the poverty line. (If the Distributive Partition fails on this poverty-line view, then it should fail *a fortiori* on more robust distributive views, such as egalitarianism or maximin.) Imagine that O_1 is welfare maximizing relative to O_2 , but that O_1 is worse with respect to poverty (for simplicity, because some persons are below the poverty line in O_1 but no one is below the line in O_2 .) Why think that a regulatory agency should ignore egalitarian considerations and choose O_1 ?

One possibility is that the legislature will *directly reverse* regulatory choices that are welfare maximizing, but, all things considered, morally bad by virtue of their distributive defects. In this example, the agency will choose O_1 over O_2 and the legislature will then promulgate a statute requiring O_2 . This version of the Distributive Parti-

201. This view is also defended by PHILIPPE VAN PARIJS, *REAL FREEDOM FOR ALL: WHAT (IF ANYTHING) CAN JUSTIFY CAPITALISM?* (1995).

202. See, e.g., KAPLOW & SHAVELL, *supra* note 37; Aanund Hylland & Richard Zeckhauser, *Distributional Objectives Should Affect Taxes but Not Program Choice or Design*, 81 SCAND. J. ECON. 264 (1979); Polinsky, *supra* note 41, at 119-27; Steven Shavell, *A Note on Efficiency vs. Distributional Equity in Legal Rulemaking: Should Distributional Equity Matter Given Optimal Income Taxation?*, 71 AM. ECON. REV. 414 (1981). Chris Sanchirico summarizes the claim:

The New [Efficiency] Rationale . . . shunt[s] distributional concerns across disciplinary boundaries from private law to tax. While conditionally conceding the importance of equity as a political-philosophical value, proponents of the New Rationale insist that the legal sphere is not the proper place to pursue distributional objectives. Efforts to decrease economic inequality, it is argued, should be corralled into the tax code, while legal rules should be left in a pristinely efficient state, unsullied by distributive justice.

Sanchirico, *supra* note 13 (manuscript at 2).

tion is overly optimistic about the incentives and capabilities of legislatures. In practice, agencies have lots of “slack,” including the ability to pursue policies that increase poverty.

A second possibility is that the defects of O_1 can be cured through the income tax system by either the legislature or a tax agency. Here, the problem is that some kinds of “poverty” (properly understood) are incurable. If O_1 involves some premature deaths, such that the lifetime welfare of some dying persons is very low—sufficiently low to bring them below the level of lifetime welfare equaling “poverty”—no (future) money payments to the estates of these persons will repair their poverty. A welfare-maximizing regulatory policy might cause the death of a one-year-old toddler; the total well-being the toddler realizes over the course of his life will be drastically lower than the total well-being realized by most persons, and thus the choice of O_1 will make the toddler incurably poor.

To sum up, welfarism allows that regulators may need to attend to features of regulatory options other than their effect upon overall well-being. Which features are those? That depends both on the criteria apart from aggregate welfare (such as deontological, perfectionist, or distributive criteria) that may bear upon the moral status of governmental actions, *and* upon the institutional arrangements or “partitions” that may make different governmental bodies responsible for implementing different criteria. In this Section, I have speculated in a very preliminary way about the nature of the additional criteria and relevant arrangements. My very tentative suggestion is that regulators will *not* be able to adopt welfare maximization as their sole aim. Morality will be too complex, and institutional possibilities too crude, to permit that. But this is really just a guess on my part. The speculations in this section are simply intended to identify areas for future research, and they do not begin to approach a full analysis of the problems described here.

D. Welfarism and Neoclassicism: Some Differences

Welfarism is refurbished neoclassicism. The welfarist, like the neoclassicist, gives central importance to human welfare. Like the neoclassicist, the welfarist accords preferences a prime role in her account of well-being. Like the neoclassicist, the welfarist is focussed upon regulatory *outcomes* rather than the procedures that produce these outcomes. The neoclassicist need not take regulatory procedures to be intrinsically valuable, and the welfarist, by definition, does not. Finally, both welfarism and neoclassicism acknowledge that the moral criterion essential to these views—overall well-being in the case of welfarism, Kaldor-Hicks efficiency in the case of neoclassicism—may be supplemented by other criteria, such as deontological, perfectionist, or distributive criteria.

Nonetheless, there remain noteworthy differences between welfarism and neoclassicism, which I will briefly describe in this Part. These differences have two sources. The first source is the difference between overall well-being and efficiency. The second source is the difference between the simple preference-based view of welfare incorporated into neoclassicism and the more nuanced view that is built into welfarism, namely sophisticated preferentialism. The shift from neoclassicism to welfarism involves material changes in, albeit not fundamental alteration of, our normative theory of regulation.

1. *Efficient Welfare Losses and Inefficient Welfare Gains*

As I have already explained, an option can be Kaldor-Hicks efficient relative to an alternative but decrease overall well-being. Reciprocally, the option can be Kaldor-Hicks inefficient but increase overall well-being. The option of permitting a factory to emit a particulate that causes respiratory distress to poor persons living nearby, but that also produces a beautiful sunset viewed by rich persons from afar, is welfare decreasing. Nevertheless, this option will be Kaldor-Hicks efficient if the cumulative dollar amount that the rich persons are willing to pay for the sunset is larger than the cumulative dollar amount that the poor persons would demand in compensation for the respiratory distress.²⁰³ Or consider the option of building a road through the property of an ascetic who is barely bothered by it (since she is an ascetic), but who would still be willing to pay a lot not to have the road built (since it has *comparatively* more welfare significance for her than other things, which matter even less). This option will be welfare increasing, but it could be Kaldor-Hicks inefficient if the ascetic's willingness to pay not to have the road is large enough.

As the ascetic's case shows, the deviation between Kaldor-Hicks efficiency and overall well-being is not simply a matter of wealth effects.²⁰⁴ Given two persons (P_1 and P_2) and a regulatory choice that has the same welfare effect on the two in terms of units of overall welfare, the two persons may have different monetary equivalents for that welfare change for any of the following reasons: (1) P_1 and P_2 have roughly identical functions, mapping wealth and other features of the world onto welfare, but P_1 and P_2 are at different wealth levels (wealth effects); (2) P_1 and P_2 have roughly identical functions, mapping wealth and other features of the world onto welfare, and have

203. See *supra* Part II.B.1. Technically, the cost-benefit criterion (sum of willingness-to-pay criterion) and Kaldor-Hicks efficiency are not equivalent. The so-called "Boadway paradox" shows that cost-benefit analysis and Kaldor-Hicks efficiency can deviate in special cases. See Adler & Posner, *supra* note 28, at 190-91 & n.68. For simplicity, I here use the cost-benefit criterion and the Kaldor-Hicks criterion interchangeably.

204. See Adler, *supra* note 63, at 1398-1401.

the same wealth level, but they are differently situated with respect to other features; or (3) P_1 and P_2 have significantly *different* functions, mapping wealth and other features of the world onto welfare. Nor is it clear to me that wealth effects will, in practice, produce the largest overall divergence between efficiency and welfare. On the other hand, efficiency-type analysis, by regulatory agencies and scholars, is most readily amended to deal with the problem of wealth effects. The willingness-to-pay and -accept amounts can be weighted by a factor inversely proportional to the wealth of the person involved.²⁰⁵ Therefore my discussion here of the divergence between efficiency and welfare will focus on those kinds of effects.

Consider a regulatory policy that affects various persons who differ significantly in wealth level. The policy could be a pricing system for a regulated service that is sold both to richer and to poorer consumers (cross-subsidization of the poorer consumers will benefit them at the expense of the richer consumers, who would be better off if marginal-cost pricing were used instead). Or the policy could be a conservation policy with respect to wilderness areas and other environmental benefits that are mainly enjoyed by richer persons, where the costs of conservation would be borne by all taxpayers; or a workplace safety policy that benefits poorer workers at the expense of richer shareholders; or an antitrust policy that benefits poorer shareholders at the expense of richer consumers; or an "environmental justice" policy that locates a waste dump or some other "locally undesirable land use" in a rich neighborhood rather than a poor one. Assume that the policy is inefficient. For simplicity, assume that the policy benefits poorer persons by an amount $\$P$ and harms richer persons by an amount $\$R$ where $\$R$ is greater than $\$P$, relative to an efficient alternative. The neoclassicist will say the following about the policy:

The tradeoff between distribution and efficiency: Although an inefficient policy whose costs and benefits are borne by persons at different wealth levels may be a good thing with respect to distributive goals, it is a bad thing with respect to at least one normative goal, namely efficiency.²⁰⁶

According to the neoclassical view, if we are faced with the choice between an efficient policy and an inefficient policy that benefits poorer persons by some amount $\$P$ too small to outweigh the losses $\$R$ to richer persons, then we must balance the policy gain in one dimen-

205. See sources cited *supra* note 36 (discussing possible modification of cost-benefit analysis through incorporation of distributive weights); see also Adler & Posner, *Implementing Cost-Benefit Analysis When Preferences Are Distorted*, 29 J. LEGAL STUD. 1105 (2000) (same).

206. See, e.g., KAHN, *supra* note 11, at 189-93 (analyzing cross-subsidization in this manner).

sion (distributive goals) against the policy setback in another dimension (efficiency goals).

By contrast, according to the welfarist view, it is an open question whether the efficient policy is a good thing *in any respect*. First, the inefficient policy may be better on the distributive dimension (as the neoclassicist recognizes). Second, the inefficient policy may also be better *from the point of view of overall welfare*. Although $\$P$ is smaller than $\$R$, the richer persons' marginal utility of wealth may be sufficiently smaller than the poorer persons', such that the $\$P$ shift to the poor causes a larger effect on overall well-being than the $\$R$ shift away from the rich. For the welfarist, an inefficient policy that secures distributive gains may be an unalloyed good, while for the neoclassicist it cannot be. Even if the inefficient policy is also welfare decreasing (assuming a large enough difference between $\$R$ and $\$P$, or a small enough difference in wealth levels between richer and poorer), the policy's negative effect on overall well-being will comprise a less substantial counterweight to distributive goals for the welfarist than its inefficiency will comprise for the neoclassicist. All this has practical, not just conceptual, significance, since it means that the welfarist will end up choosing some inefficient policies that the neoclassicist would, on balance, reject.

The neoclassicist has another string to her bow. The inefficient, but egalitarian, regulatory policy and the efficient, but inequalitarian, policy are not the only options that should be considered. A third and better option might be combining the efficient regulatory policy with income tax payments and levies designed to achieve the distributive goals. Neoclassicists standardly claim the following, or something like it.

*Regulation Is Less Efficient than Taxation in the Pursuit of Distributive Goals*²⁰⁷

Given an inefficient regulatory policy, there exists an efficient alternative that, combined with suitable tax payments and levies, is just as good with respect to distributive goals and better with respect to efficiency.

I have already expressed doubt that, with respect to distributive goals, the income tax system is always as good as regulation. An efficient regulatory policy may have distributive deficits that no scheme of tax levies and payments can repair. Reciprocally, some economists have doubted whether—given an inefficient regulatory policy and the most efficient tax scheme that achieves all the distributive goals of the policy—such a tax scheme will always be more efficient than the

207. See sources cited *supra* note 202 (arguing that the tax system, not legal rules, should be used to pursue distributional goals).

policy, since there are *efficiency* deficits (specifically, work disincentives) associated with income taxation.²⁰⁸ But let us place these doubts to one side. It seems plausible that the above-stated neoclassical claim is widely, if not universally, true—that in a significant number of cases, the tax system can achieve all the distributive goals of regulation, and in a more efficient manner.

Does this eliminate the divergence between welfarism and neoclassicism? Hardly. Imagine the following: the inefficient policy causes a gain of $\$P$ for the poorer persons, which is smaller than the loss of $\$R$ incurred by the rich. Our distributive goal is to bring all persons above the poverty level. A tax scheme combined with the efficient regulatory alternative would give $\$P$ to the poorer persons, which makes it just as good as the inefficient policy with respect to this distributive goal, and would cost the richer persons $\$T_r$. The tax scheme would also cost other Taxpayers $\$T_o$. $\$T_r$ plus $\$T_o$ is less than $\$R$, so the scheme is more efficient than the inefficient policy.

	Poor	Rich	Other Taxpayers
Efficient Policy	—	—	—
Inefficient Policy	$\$P$	$-\$R$	—
Tax Scheme + Efficient Policy	$\$P$	$-\$T_r$	$-\$T_o$

How do the tax scheme and the inefficient policy compare within neoclassicism and welfarism? Now it is the neoclassicist who perceives an unalloyed good! For the neoclassicist, the tax scheme is unqualifiedly better than the inefficient policy: It is more efficient, and it is no worse with respect to distributive goals. By contrast, the welfarist agrees that the scheme is just as good with respect to distributive goals, but she notes that it may be *worse* with respect to welfare. Imagine that the taxpayers who incur $\$T_o$ are poorer than the richer persons who benefit from the difference between $\$R$ and $\$T_r$. Then, the tax scheme might be welfare decreasing even though $\$T_o$ is less than $\$R$ minus $\$T_r$. If so, the welfarist would need to balance the overall-welfare goal against the distributive goal—and, in this case, would presumably end up rejecting the tax scheme on overall-welfare grounds, while the neoclassicist would have no reason whatsoever to reject the tax scheme.

It might be objected that the kind of tax scheme envisioned by the neoclassicist is a tailored arrangement that taxes the richer persons

208. See Sanchirico, *supra* note 13.

(those harmed by the inefficient policy) for the benefit of the poorer persons (those helped by it), while leaving other taxpayers unaffected.

	Poor	Rich	Other Taxpayers
Efficient Policy	—	—	—
Inefficient Policy	\$P	-\$R	—
Tax Scheme + Efficient Policy	\$P	-\$P	—

Indeed, the welfarist, like the neoclassicist, *does* view this tailored scheme as an unalloyed improvement over the inefficient policy. However, that is only because the scheme is Pareto-optimal, relative to the policy. The notion that the taxing authorities will duplicate the distributive benefits of inefficient regulation through tailored, Pareto-optimal schemes that carefully burden only those who would be burdened by the regulation is highly unrealistic. We live in a world where government officials must generally choose between Pareto-noncomparable options. This kind of choice, *inter alia*, is one that permits welfare and efficiency to diverge.

2. Restricted and Unrestricted Preferences

Neoclassicism incorporates an unrestricted preference-based theory of welfare. According to this view, P is comparatively benefited by outcome S_1 , relative to S_2 , just in case P prefers S_1 to S_2 . As I have already explained, the difficulty with this view is that P 's preference is simply a ranking that is conceptually linked, in some way, with P 's actions.²⁰⁹ (The conceptual linkage goes something like this: P prefers S_1 to S_2 just in case, if P believes that action A_1 leads to S_1 and action A_2 leads to S_2 , P chooses A_1 .) However, a wide range of factors, unconnected to P 's welfare, can motivate his outcome rankings and, therewith, his actions. P can prefer S_1 to S_2 on moral grounds, because he thinks S_1 is required by some official or other role that he holds, or because he believes S_1 is in the interests of some community to which he belongs and which he feels himself obliged to support. Sophisticated preferentialism, in contrast to neoclassicism, posits that only a subset of P 's preferences—"restricted" preferences—are welfare productive for P . Think of the restriction idea this way:

209. See *supra* Part II.C.

What Are Restricted Preferences?:

In some cases, (1) P prefers S_i to S_j , and (2) P would prefer S_i to S_j under idealized conditions, but S_i is not better for P 's welfare than S_j . In other cases, (1) P prefers S_i to S_j , and (2) P would prefer S_i to S_j under idealized conditions, and S_i is better for P 's welfare than S_j . "Restricted" preference is that property of P , S_i , and S_j , such that if the property obtains and if (1) P prefers S_i to S_j , and (2) P would prefer S_i to S_j under idealized conditions, then S_i benefits P .

This is hardly a full analysis of the concept of restricted preference; I do not yet have such an analysis to offer. Here are some possibilities: (a) P 's preference for S_1 over S_2 is restricted if the ground for P 's preference is *not* his belief that S_1 is morally required or required by some legal or social norm; (b) P 's preference for S_1 over S_2 is restricted if P prefers S_1 under some description that essentially involves P 's own existence; (c) P 's preference for S_1 over S_2 is restricted if the ground for P 's preference is P 's belief that S_1 improves his welfare; or (d) P 's preference for S_1 over S_2 is restricted if S_1 is better for P in light of some objective welfare good. Each of these possibilities has flaws, and I am not sure how to balance the flaws against the merits.

Sophisticated preferentialism is a class of welfare-accounts that, one way or another, restrict preferences. Welfarism is a family of moral views that give *prima facie* weight to overall well-being and incorporate *one or another members* of the sophisticated-preferentialist class of welfare views. In this Article, I have simply argued that the right moral view is some member of the welfarist family of moral views, without making further claims as between family members. On the other hand, it can be said that *any* member of the welfarist family of moral views—because that view restricts preferences, in some way—will diverge from neoclassicism.

Here is a simple way to see the divergence. An agency must choose between S_1 and S_2 . Some persons prefer S_1 to S_2 , or S_2 to S_1 , but they do not restrictedly prefer S_1 to S_2 or S_2 to S_1 . Call these persons disinterested. Other persons prefer S_1 to S_2 , or S_2 to S_1 , and they further satisfy the sophisticated preferentialist's criteria for welfare—namely, these preferences are restricted and survive idealization. Call these the interested persons. The neoclassicist will consider the effect of the S_1/S_2 choice, both on the interested persons and on the disinterested persons, in determining which option maximizes efficiency. By contrast, the welfarist will focus solely on the interested persons. Changes in the number of disinterested persons, or in the intensity of their preferences, will not change the welfarist's evaluation of the S_1/S_2 choice with respect to overall well-being.

For example, the S_1/S_2 choice might be the choice of an antitrust agency: S_1 means having a cartel in a particular market, while S_2

means having competitive conditions in that market. Then, plausibly, consumers and firm shareholders are interested persons, while the citizen who prefers or disprefers S_1 on purely moral grounds (perhaps because he thinks that cartels prevent destructive competition, or conversely because he thinks they diminish consumer welfare) is disinterested. The welfarist will ignore the disinterested citizen, while the neoclassicist should not—at least if the neoclassicist is consistent about sticking to the unrestricted preference-based view of well-being that is standard within welfare economics. After all, the citizen does prefer S_1 to S_2 or S_2 to S_1 ; this preference may well satisfy all the standard axioms of expected-utility theory (it may well be complete, reflexive, and so on); it can be assumed to have the requisite linkage to the citizen's actions; and it can give rise to some monetary willingness to pay for S_1 over S_2 or S_2 over S_1 .

Yet it is counterintuitive, unnecessary, and maybe downright incoherent to think that the existence of a citizen with moral views for or against cartelization should itself change the moral status of cartelization—in particular, the effect of cartelization on overall well-being (or, for that matter, on efficiency).²¹⁰ This is what leads the welfarist to insist on preference restriction, and to train his attention solely on interested persons (here, consumers and shareholders) in calculating aggregate welfare. Similarly, suppose S_1/S_2 is the choice facing an occupational health and safety agency, with S_1 requiring firms to impose certain workplace safety devices and S_2 permitting them to omit such devices. Then the preferences of labor activists, public-interest lawyers, or left-leaning scholars who are morally motivated to give a higher ranking to S_1 , as well as the preferences of free-market ideologues who are morally motivated to give a higher ranking to S_2 , would be ignored by the welfarist in evaluating the S_1/S_2 choice with respect to overall well-being. By contrast, these preferences would *not* be ignored by the neoclassicist in evaluating that choice with respect to efficiency.

In the above cases, I have chosen actors whose preferences are purely morally motivated to fill the role of disinterested persons. I took this tack because a purely morally motivated preference is the hallmark of a preference that is *not* restricted and that would not be counted as welfare productive by any sophisticated preferentialist. Whatever the correct account of restriction, it must, at a minimum, exclude the case where P prefers S_1 just because he thinks S_1 is morally required (and would otherwise prefer S_2 , or be indifferent) as a case where S_1 improves P 's welfare. However, as I have suggested, the correct account of restriction may end up describing other kinds

210. See generally Adler & Posner, *supra* note 205 (arguing that agencies should ignore moral and other disinterested preferences when performing cost-benefit analysis); Railton, *supra* note 35, at 55, 71-72 (same).

of persons as disinterested—for example, persons motivated by legal obligations, social norms, community loyalties, or considerations which they do not take to be self-regarding. If so, the divergence between neoclassicism and welfarism will be yet broader.

In practice, it seems, regulatory agencies generally *deviate* from the unrestricted preference-based view of welfare and implicitly ignore persons who have policy preferences but seem to be disinterested.²¹¹ The Federal Trade Commission and the Department of Justice do not incorporate citizen preferences for and against antitrust policies into their cost-benefit analyses of these policies. Otherwise, the standard regulatory view that cartels and monopolies are generally inefficient would be harder to support, because sufficiently strong pro-cartel or pro-monopoly preferences on the part of citizens plus firm profits could outweigh the consumers' loss. Similarly, the Occupational Safety and Health Administration (OSHA) looks to workers, consumers, and maybe shareholders, but not to the entire universe of people who have preferences over OSHA's policies.

The one exception is environmental law, where the Environmental Protection Agency and other environmental agencies have begun computing "existence values" for environmental amenities. Individuals who will never physically interact with a particular endangered species, wilderness area, or pristine bit of the atmosphere are asked how much they would be willing to pay for that amenity to remain in place. These values are then factored into agency assessment of environmental policies.²¹² The neoclassicist would say that the environmental agencies' unique attention to disinterested preferences in determining the welfare or efficiency effect of policies is appropriate, and that other regulatory agencies should follow the environmental agencies' lead. The welfarist would say that the environmental agencies' unique attention to disinterested preferences is deeply misguided—that existence values should no more be calculated for environmental amenities than for antitrust policies, workplace policies, transportation policies, redistributive policies, housing policies, military policies, and so on, and that it is the environmental agencies whose practices ought to change.

211. See Adler & Posner, *supra* note 205 (surveying agency practice).

212. See *id.* (critically discussing EPA's use of existence values); CONTINGENT VALUATION: A CRITICAL ASSESSMENT (Jerry A. Hausman ed., 1993) (same); MARK SAGOFF, THE ECONOMY OF THE EARTH (1988) (same); Symposium, *Contingent Valuation*, J. ECON. PERSP., Autumn 1994, at 3-64 (same).

3. *Incomparability and Precise Equality*

Options can be neither better nor worse (NBNW) with respect to overall well-being.²¹³ One option is NBNW than another with respect to overall well-being if (a) the two options are equally good, or (b) the two options are incomparably good. If two options (S_1 and S_2) are equally good with respect to overall well-being, then any third option is better than S_1 just in case it is better than S_2 and, conversely, worse than S_1 just in case it is worse than S_2 . Incomparability is less demanding: if S_1 and S_2 are incomparably good with respect to overall well-being, then a third option can be better than S_1 without being better than S_2 , or worse than S_1 without being worse than S_2 .

In practice (I would guess) agencies are not often faced with a choice between options that are equally good with respect to overall well-being. Why? Where NBNW options are good for overall welfare in different ways—where, for example, the extent of physical pleasure is greater in S_1 , while the extent of friendship is greater in S_2 , or longevity is greater in S_1 , but convenience is greater in S_2 —then there will virtually always be a hypothetical option which is a small improvement over one option but still NBNW than the other. If so, the two options by definition are not precisely equal.²¹⁴

On the other hand, it could well be the case that regulatory agencies *do* frequently confront incomparable options. Recall my account of interpersonal comparisons: S_1 is better than S_2 with respect to overall well-being, or worse than S_2 with respect to overall well-being, if and only if all idealized observers would have convergent restricted preferences as between S_1 and S_2 . Will this occur if S_1 is somewhat, but not dramatically, better than S_2 with respect to one welfare good while S_2 is somewhat, but not dramatically, better than S_1 with respect to another? (For example, what if there are 1000 more annual deaths in S_1 , but the unemployment rate in S_2 is a half percentage point higher? What if consumer surplus in S_1 is \$100 million greater than in S_2 , but twenty endangered species go extinct? What if workplaces in S_1 are moderately safer, but the quality of education is moderately worse?) When we think hard about the welfare tradeoff between such options, we may conclude that the options are welfare-comparable—that one option does emerge as better, and the other as worse. But it is at least plausible to think that a significant fraction of agency decisions involve incomparable, and thereby NBNW, alternatives.

213. See Ruth Chang, *Introduction to INCOMMENSURABILITY, INCOMPARABILITY AND PRACTICAL REASON*, *supra* note 163, at 1-34 (defining and discussing incomparability); sources cited *supra* note 163 (same).

214. See Ruth Chang, *Introduction to INCOMMENSURABILITY, INCOMPARABILITY AND PRACTICAL REASON*, *supra* note 163, at 23-27 (concluding that the “small improvement” argument is a plausible argument for incomparability).

Would this create a divergence with neoclassicism? Note that two options can be NBNW with respect to efficiency. Clearly, two options can be equally efficient; it also turns out that two options can be incomparably efficient.²¹⁵ On the other hand, the proportion of regulatory agency choices that are either equally efficient or incomparably efficient seems to be low.²¹⁶ So it is at least plausible that many agency choices are NBNW with respect to overall welfare, but few agency choices are NBNW with respect to efficiency.

Imagine that two options are NBNW with respect to a particular moral criterion X^* (where X^* is the criterion of overall well-being, for the welfarist; and the criterion of efficiency, for the neoclassicist). How should the agency choose between the options? There may be some other criteria, $X_1, X_2 \dots X_n$, that possess moral force, and that the agency is charged with implementing. (That is, there may be other criteria not covered by a "partition" that effectively places them within the jurisdiction of legislatures, courts, or the taxing authorities, and outside the jurisdiction of regulatory agencies). If so, the agency should choose the option that is better with respect to the applicable $X_1, X_2 \dots X_n$.

But it may be that no such $X_1, X_2 \dots X_n$ exist; or (more likely), they may exist, but the option may also be NBNW with respect to the $X_1, X_2 \dots X_n$. In such a case, I suggest, the agency is free to choose at random between the options. It possesses *moral discretion*. Where an agency's alternatives are NBNW with respect to the applicable moral criteria $X^*, X_1 \dots X_n$, the agency might as well decide between the alternatives by flipping a coin. This is clearly true when the options are equally good with respect to the $X^*, X_1 \dots X_n$; and it is also true (I have argued elsewhere) when the options are incomparably good with respect to the $X^*, X_1 \dots X_n$.²¹⁷ If, for example, our beneficial me-

215. This is a result of the Scitovsky paradox. Take two outcomes O_1 and O_2 that involve the paradox, that is, there is some redistribution from Winners to Losers in O_1 that makes it Pareto-superior to O_2 , but there is also some redistribution from Winners to Losers in O_2 that makes it Pareto-superior to O_1 . Then, O_1 and O_2 are NBNW with respect to efficiency. However, they are not precisely equal with respect to efficiency, since there will typically be at least some O^* that is efficient relative to O_1 , but not O_2 (or vice versa). For example, create O^* by taking each person's holding of each good in O_1 and increasing it slightly. That guarantees that O^* is efficient relative to O_1 , but there could still be some redistribution in O_2 that makes it Pareto-superior to O^* .

216. Why do I say this? Cost-benefit analysis is closely related, if not equivalent, to Kaldor-Hicks efficiency, and it is highly unusual for an agency performing cost-benefit analysis to conclude that one option has neither positive nor negative net benefits relative to another.

217. See Adler, *supra* note 63, at 1401-08. Amartya Sen makes this very claim:

Some see completeness as a necessary requirement of consequential evaluation, but it is, of course, nothing of the sort. A consequentialist approach does involve the use of maximizing logic . . . [but] [m]aximization only requires that we do not choose an alternative that is worse than another that can be chosen instead. If we cannot compare and rank two alternatives, then choosing either from that pair will fully satisfy the requirement of maximization.

teor can be used to help Phil's or Pat's emphysema, and the choice of one or the other would not affect the fair distribution of welfare (which would be true if, for example, poverty-line egalitarianism obtained and both Phil and Pat were well above the poverty line) or perfectionist values, then the meteor-possessing agency is morally free to give the rock to either Phil or Pat. Either choice is morally permitted; neither is morally required or morally prohibited.

The upshot of this analysis is that welfarism may confer greater moral discretion upon agencies than neoclassicism. If, in practice, agency choices are rarely NBNW with respect to efficiency, then agencies rarely have moral discretion (within the framework of neoclassicism).²¹⁸ If, in practice, agency choices are frequently NBNW with respect to welfare, then agencies may frequently have moral discretion (within the framework of welfarism), depending upon the applicability and content of perfectionist, distributive, deontological, and other such supplementary moral criteria.

4. *The Possibility of Paternalism?*

Let me distinguish between weak and strong paternalism.²¹⁹ Strong paternalism holds that persons can be made better off in the teeth of their preferences. That is, (1) *P* prefers S_1 to S_2 , but nonetheless (2) S_2 can be better for *P*'s welfare than S_1 . Weak paternalism holds that persons can be no better off notwithstanding the satisfaction of their preferences. That is, (1) *P* prefers S_1 to S_2 , but nonetheless (2) S_1 can be no better for *P*'s welfare than S_2 .²²⁰

Neoclassicism denies both strong paternalism and weak paternalism. This is easy to see. Neoclassicism incorporates an unrestricted preference-based view of well-being, which says that *P* is better off with S_1 , as compared to S_2 , just in case he prefers S_1 to S_2 . If *P* does prefer S_1 to S_2 , then—the neoclassicist concludes—he is better off with S_1 , not S_2 (*pace* strong paternalism). The neoclassicist also concludes that this is a real welfare improvement, rather than a case where S_1 is not better than S_2 for *P* (*pace* weak paternalism).

Sen, *The Discipline of Cost-Benefit Analysis*, *supra* note 40, at 940.

218. This is overstated: an outcome could be better than another with respect to efficiency, while the second outcome is better with respect to $X_1, X_2 \dots X_n$, such that the outcomes are NBNW with respect to the balance of $X^*, X_1 \dots X_n$ and the agency has moral discretion on balance. Nonetheless, it seems correct to say that if choices are rarely NBNW with respect to X^* , it is less likely that they will be NBNW with respect to $X^*, X_1 \dots X_n$.

219. For discussions of paternalism, see GERALD DWORKIN, *THE THEORY AND PRACTICE OF AUTONOMY* (1988); JOHN KULTGEN, *AUTONOMY AND INTERVENTION: PARENTALISM IN THE CARING LIFE* (1995); DONALD VANDEVEER, *PATERNALISTIC INTERVENTION* (1986); and Joel Feinberg, *Legal Paternalism*, 1 *CAN. J. PHIL.* 105 (1971).

220. See Feinberg, *supra* note 219 (distinguishing between "weak" and "strong" paternalism). My distinction is quite different.

Welfarism denies strong paternalism, but it admits weak paternalism. This is pretty easy to see, too. Welfarism incorporates a sophisticated preferentialist view of well-being. This says (placing to one side, for now, the issue of restriction) that *P* is better off with *S*₁, as compared to *S*₂, just in case (1) *P* prefers *S*₁, and (2) *P* would prefer *S*₁ under suitably idealized conditions. If *P* does prefer *S*₁ to *S*₂, then—even if he would prefer *S*₂ under suitably idealized conditions—he does *not* actually prefer *S*₂ to *S*₁, and he therefore cannot be made better off with *S*₂. This shows the inconsistency between sophisticated preferentialism and strong paternalism. On the other hand, if *P* prefers *S*₁ to *S*₂ but would prefer *S*₂ under idealized conditions, then *P* is not made better off by *S*₁. This shows how weak paternalism flows from sophisticated preferentialism.

An example may make the point clearer. *P* prefers sitcoms to opera but would prefer opera under suitably idealized conditions. The public television station's decision to show opera instead of sitcoms does not make *P* better off, because he actually prefers sitcoms. So the strong paternalist view that *P* is a *beneficiary* of the station's choice of opera is a mistake. Yet the fact that *P* actually prefers sitcoms does not mean that he benefits if, instead of opera, the station chooses to broadcast sitcoms. After all, *P*'s idealized preference is in favor of opera. Sophisticated preferentialism gives equal weight to this idealized preference, along with *P*'s actual preference, in determining what improves his welfare. For the sophisticated preferentialist, *neither* programming choice is better for *P*'s welfare. *P* is not better off with the sitcom, nor is he better off with the opera, since the dual and conjunctive conditions for a welfare benefit—actual preference satisfaction plus idealized preference satisfaction—are not true of either option. Among other things, this dual and conjunctive account of welfare bears out weak paternalism.

What does this mean for regulatory agency policy? It means that the agency can (weakly) decide that some persons who believe themselves affected by an agency's choice are actually unaffected. Persons who are counted as "Winners" or "Losers" under neoclassicism may become "Neutrals," neither benefited nor harmed by the choice, and therefore ignored for purposes of evaluating that choice in light of overall well-being. Consider, as a plausible example, the laws prohibiting recreational use of drugs.²²¹ Persons who prefer drug use must be counted as Losers from an anti-drug policy within a neoclassical framework,²²² but they may well be seen as Neutrals by the welfarist

221. See generally DOUGLAS N. HUSAK, *DRUGS AND RIGHTS* (1992) (presenting detailed philosophical argument against prohibition of recreational drug use).

222. More precisely, *certain* persons who prefer drug use must be counted as Losers from an antidrug policy, within a neoclassical framework. It is arguably consistent with that framework to say that drug addicts (persons whose preference for drug use is effec-

(if she thinks they would disprefer drug use under idealized conditions).

It also may mean, a bit more robustly, that where actual preferences are changing, agencies can rely upon idealized preferences as the sole criterion for determining an option's welfare effect. Imagine that P now prefers S_1 to S_2 , and that he will continue to prefer S_1 if it is chosen by the agency; but he would change his mind and come to prefer S_2 if *that* were chosen. Then, neoclassicism has no clear basis for determining which option benefits P . Here, it is neoclassicism which may well end up classifying P as a Neutral, while the welfarist could plausibly say that whatever option P would ideally prefer is the one that benefits him.²²³ In such a case (on a plausible specification of sophisticated preferentialism), P is a Winner with S_1 if he ideally prefers S_1 , and P is a Winner with S_2 if he ideally prefers S_2 .

E. Welfarism and Proceduralism: A Large Difference

The difference between welfarism and proceduralism is both larger and simpler than that between welfarism and neoclassicism. A proceduralist theory, by definition, accords intrinsic moral significance to the decisionmaking procedures that regulators follow. A welfarist theory, by definition, does not. Imagine that a regulatory agency, or an executive or legislative body with authority over the agency, is making a threshold determination about which procedures the agency ought to employ. For simplicity, imagine that the regulatory agency has a large set of substantive options that are legally available to it (specifically, the options of issuing various legal directives, plus the option of issuing no directive), and that the threshold or procedural choice is a binary choice between P_i and P_j , where P_i and P_j are alternative possible routines that the agency could follow in choosing a substantive option. P_i could be cost-benefit analysis, regulatory negotiation, civic-republican deliberation, a familiar notice-and-comment process followed by an intuitive judgment on the regulator's part, or any other combination of standard or esoteric actions by the agency that would result in a particular substantive option. P_j could be any such procedure distinct from P_i . Then the welfarist says this: Whether the agency or oversight body should choose P_i or P_j is wholly a function of which substantive options the two proce-

tively coerced because their preference is grounded in the withdrawal symptoms and other suffering they anticipate if they stop using drugs) and persons whose preferences are internally conflicted are *not* Losers from an antidrug policy. However, the neoclassicist cannot explain why the nonaddict who (without internal conflict) just likes drugs is not a Loser.

223. See Adler & Posner, *supra* note 205 (suggesting that agencies could look to fully informed preferences where actual preferences are changing).

dures can be expected to produce.²²⁴ Relatedly, if P_i and P_j can be expected to eventuate in the very same option, or in the very same slate of options with the very same probability assigned to each option, then P_i and P_j are morally identical. It would, in that case, be a matter of indifference which procedure is chosen at the threshold.

In short, for the welfarist, different procedures P_i and P_j should be understood as *lotteries* over substantive options (lotteries over possible regulatory directives). Welfarism provides a set of substantive criteria (W^* , X_1 , X_2 . . . X_n , where W^* is overall well-being and each X_i is another moral criterion) for evaluating the substantive options. When fully specified, welfarism should also provide a theory of moral choice under uncertainty. That theory will specify how regulators or oversight bodies should choose between actions whose outcomes are uncertain, as a function of the criteria (W^* , X_1 , X_2 . . . X_n) plus the probability of each outcome that the action may produce. The threshold procedural choice between P_i and P_j is a matter of predicting which substantive options (directives) the different procedures may produce, with which probabilities, and then applying the substantive criteria (W^* , X_1 , X_2 . . . X_n) together with the welfarist's theory of choice under uncertainty. By contrast, the proceduralist claims the existence of some procedural criterion X_P such that two procedures can differ with respect to X_P , even though they are expected to produce the very same option (or the very same probabilistic slate of options). In the proceduralist's eyes, the threshold choice between P_i and P_j should then depend, wholly or at least partly, on how P_i and P_j fare with respect to X_P .

Proceduralism consists of a class of theories, each of which specifies X_P in different ways. The civic-republican proceduralist thinks that X_P favors civic-republican deliberation. The pluralist proceduralist thinks that X_P favors a process of interest bargaining between groups. The collaborative-governance proceduralist thinks that X_P favors a certain kind of collaboration between affected parties. Each such theory gives substantial, if not conclusive, weight to the specified X_P in determining procedural choices. The proceduralist does not (or need not) deny that outcomes can be evaluated, apart from the procedures that produce them. The proceduralist simply thinks that the threshold choice of procedures does not *reduce* to the evaluation of outcomes. So, given any welfarist theory (W^* , X_1 , X_2 . . . X_n), that theory will be paired with a matching family of proceduralist theories (W^* , X_1 , X_2 . . . X_n , X_P) that specify X_P in different ways. The shift to a given welfarist theory from some proceduralist theory within the

224. I am ignoring here (1) the direct costs of the two procedures, which may differ, and (2) the indirect effect that this procedural choice may have on other choices, for example, future procedural choices. For simplicity, and to make the contrast with proceduralism as clear as possible, I have placed these issues to one side.

matching family simply involves *abandoning* any bedrock moral role for the favored procedural type X_P . X_P no longer constitutes a basic moral criterion. That shift has no implications for a choice between substantive regulatory options, but it clearly may have large implications for the threshold choice of procedures by regulators or by the executive or legislative bodies that oversee them.

All of this is very abstract. What, concretely, are the procedural implications of welfarism? Which procedures, in which context, are regulators required to adopt or to avoid? Answering these questions requires a specific view as to the substantive moral criteria (W^* , X_1 , X_2 . . . X_n) applicable to regulators. In this Article I have been diligently noncommittal about such specifics, beyond claiming that the right view includes W^* and is therefore welfarist. Assume that the *only* substantive criterion applicable to regulators is the criterion of overall well-being, either because utilitarianism is true (which seems implausible), or because other criteria obtain but have been “partitioned” away from regulators (which is less implausible, but still doubtful). In that simple case, the appropriate procedure for the regulator or oversight body to choose, given the alternatives of P_i and P_j , or any threshold set of procedural alternatives, is the *welfare-maximizing procedure*: the procedure such that the slate of outcomes it can be expected to produce is welfare maximizing, relative to the slate produced by alternative procedures.

Professor Eric Posner and I have elsewhere argued that the welfare-maximizing procedure for regulatory agencies is often, although not always, the procedure of cost-benefit analysis.²²⁵ Whether a given procedure is welfare maximizing, as against alternatives, depends on the set of substantive options to which the procedures will be applied. Posner and I claim that cost-benefit analysis is welfare maximizing, as against standardly proposed alternatives, except where the agency’s substantive options are skewed by wealth effects²²⁶—except where the Losers and Winners from some option, as against some other option, have significantly different wealth levels. *If* the Posner-Adler claim is correct, and *if* the only moral criterion applicable to regulatory choices is the aggregate-welfare criterion, then regulators would do best, morally speaking, by employing cost-benefit analysis as the basic procedure for choice between (legally available) directives and other options, to be set aside only where wealth skews are large.

Again, however, there is serious reason to doubt that overall well-being is the sole moral criterion applicable to regulatory choices. Even taking into account the role of other governmental institutions,

225. See Adler & Posner, *supra* note 28.

226. Or where cost-benefit analysis is too expensive, given the small expected difference between the options with respect to overall well-being.

the set of criteria that bear upon regulatory options and, derivatively, procedures will likely include not just the aggregate welfare criterion W^* , but also one or more criteria X_i that are distributive, deontological, or perfectionist in form. Nor is there any reason to think that cost-benefit analysis is the morally best procedure, all things considered. There is no reason to think that—comparing now the probabilistic slate of options that cost-benefit analysis can be expected to produce to the probabilistic slate of options that each alternative feasible procedure can be expected to produce—the first slate will be better in light of a fuller set of moral criteria including at least one distributive, deontological, or perfectionist X_i in addition to W^* . For example, there is no reason to think that cost-benefit analysis tracks how options fare with respect to distributive criteria, or how options fare with respect to distributive criteria, as balanced against the criterion of overall well-being. At least Posner and I see no such reason. However, we *do* make a weaker claim, which is additional to the basic assertion that cost-benefit analysis is welfare maximizing. We claim that cost-benefit analysis is *one component* of the procedure that is morally best, all things considered.²²⁷ Given whatever fuller set of welfarist criteria (W^* , X_1 , X_2 . . . X_n) is applicable to regulatory choices, the morally best procedure in light of that set will include cost-benefit analysis as a distinct part.

The idea is this: The morally best procedure for regulators in light of (W^* , X_1 , X_2 . . . X_n) will start with a serial evaluation of the options, first in light of W^* , then in light of each X_i ; and it will then move on to some kind of balancing of the different criteria in case no option is best in light of all of them. But the best way to evaluate options in light of W^* is to perform cost-benefit analysis. Cost-benefit analysis should be performed at that initial stage of the morally best procedure where agencies are attempting to determine how options fare with respect to W^* . It is not a means for determining how options fare with respect to the X_i , nor for the latter-stage balancing, if that becomes necessary.

Posner and I could well be incorrect in making the basic claim that cost-benefit analysis is welfare maximizing. In addition, even if we are correct in that claim, we could be quite wrong to advance the additional (and much more speculative) assertion that cost-benefit analysis is a component of the procedure that regulators would do best to follow, once distributive, perfectionist, and deontological criteria are brought into play alongside welfare maximization. It could be the case that regulatory negotiation, not cost-benefit analysis, is the welfare-maximizing procedure. For example, if agencies can be

227. Assuming wealth skews are not too large and cost-benefit analysis is not too expansive.

predicted to make lots of mistakes in the application of cost-benefit analysis, or to regularly distort it to their own ends, then the oversight body making a threshold choice between cost-benefit analysis and other procedures has reason to think that the slate of options produced by cost-benefit analysis will not be welfare maximizing. Or it could be the case that civic-republican deliberation, collaborative governance or for that matter pluralist interest-bargaining, is the morally best procedure once distributive, perfectionist, and deontological criteria are brought into play.

My point in this article is not to repeat or bolster the Adler-Posner case for cost-benefit analysis, but rather to bolster the implicit Adler-Posner claim about how cost-benefit analysis should be evaluated. Cost-benefit analysis should not be rejected as a regulatory or governmental procedure merely because it lacks the participatory or deliberative features favored by civic republicans, pluralists, and other proceduralists. Evaluating regulatory procedures means first specifying a substantive moral framework (one that includes at least W^*) and then making an empirical assessment of the tendency of different procedures to produce better or worse outcomes or lotteries of outcomes in light of that substantive framework. Regulators *could* still be morally required to do what the proceduralist claims, but if they are, it will be because the favored procedure tends to maximize overall welfare, distribute welfare fairly, increase overall perfection, and induce regulatory compliance with deontological norms. *That* is what those scholars who commend participatory or deliberative procedures, as against (say) cost-benefit analysis, now need to show.

V. CONCLUSION

Any plausible moral theory of regulation will give a substantial role to human welfare. What distinguishes theories, of course, is what precise role welfare plays, and how specifically that concept is analyzed. I have appropriated the term “welfarist” (perhaps unfairly) to denote a particular theory or, more exactly, a particular family of regulatory theories. This family of theories offers a mixed answer to various conceptual and normative questions crucial to moral evaluation.

A welfarist theory draws a conceptual link between preference and welfare, but falls short of equating welfare and preference-satisfaction. The theorists I have described as “neoclassicists” stipulate that one outcome benefits some person, as compared to another outcome, if and only if the person prefers the first outcome. By contrast, the welfarist asserts that an outcome must satisfy a more complicated set of conditions to be welfare-enhancing; it must be not only preferred, but ideally preferred, and both sets of preferences need to be appropriately “restricted” so as to exclude moral and other disin-

terested preferences. (This is the view of welfare I have called “sophisticated preferentialism.”²²⁸)

Welfarists also offer a mixed answer to the question, What is the moral significance of aggregate or overall well-being? Utilitarians view well-being as morally conclusive: the morally required choice, in any choice situation, is that choice whose outcomes maximize overall well-being. Other moral theories give zero weight to overall well-being; consider a wholly egalitarian theory that requires actors always to choose that option which best promotes the equality of well-being, or a “minimalist” theory that merely enjoins them to comply with certain deontological requirements and otherwise leaves actors morally unconstrained.²²⁹ By contrast with egalitarians and minimalists, welfarists claim that overall well-being has some weight in moral evaluation; by contrast with utilitarians, welfarists may also accord moral weight to other criteria, such as distributive, perfectionist, or deontological criteria.²³⁰

Welfarism, as I have defined it, does not generally specify which moral criteria, apart from aggregate welfare, possess moral weight. In other words, a moral theory is “welfarist” if it adopts a sophisticated preferentialist view of welfare and if it sets forth a group of moral criteria (W^* , X_1 , X_2 . . . X_n), where W^* is the criterion of overall well-being and the X_i are other criteria.²³¹ I have placed only two restrictions on the content of the X_i : first, that Kaldor-Hicks efficiency is not an X_i ; and second, that no X_i is a procedural criterion, a criterion in light of which regulatory outcomes identical but for the regulatory procedures that produced them can fare differently. In short, welfarists do not accord intrinsic moral significance to regulatory procedure or to the pursuit of Kaldor-Hicks efficiency.

I have defined welfarism in this way so as to differentiate welfarist theories from the two types of theories that have dominated legal scholarship about regulation: neoclassical theories and proceduralist theories. Neoclassicists equate welfare with preference satisfaction and accord moral weight to Kaldor-Hicks efficiency rather than overall well-being; proceduralists incorporate one or another procedural criterion into their moral theories. The family of welfarist theories, by definition, does not overlap either with the family of neoclassical theories or with the family of proceduralist theories.

228. See *supra* Section II.C.2.

229. See *supra* Section IV.B.

230. See *supra* Section IV.C.

231. The welfarist is not required to set forth criteria (X_1 , X_2 . . . X_n) independent of W^* . Utilitarianism, which specifies W^* as the sole moral criterion, also counts as a kind of welfarism. However, as I have explained, it seems unlikely that utilitarianism is the correct moral theory. See *supra* text accompanying note 168.

This Article has argued in favor of welfarism,²³² and against neoclassicism and proceduralism.²³³ I have also suggested that welfarism is closer in spirit to neoclassicism than it is to proceduralism.²³⁴ Welfarists, like neoclassicists, reject objectivist, hedonic, and other non-preferentialist accounts of well-being; like neoclassicists, welfarists believe that some aggregative measure, linked to welfare (be it Kaldor-Hicks efficiency or overall well-being) must play a role in moral evaluation; and, like neoclassicists, welfarists deny that governmental procedures are morally significant *per se*.²³⁵ Law-and-economics scholars and others who adopt a neoclassical view are mistaken, but not dramatically so. The moral theory of regulation proposed by law-and-economists is salvageable, rather than fundamentally flawed, and I have tried to show here what the appropriate modifications to that theory are.

232. See *supra* Sections II.C.2, IV.A-B.

233. See *supra* Parts II, III.

234. See *supra* Section IV.D-E.

235. This last statement is a bit imprecise. Given my definition of "neoclassicism," a neoclassicist could include a procedural criterion in his theory, but he is not required to do so.