IMPLEMENTING COST-BENEFIT ANALYSIS WHEN PREFERENCES ARE DISTORTED

*MATTHEW D. ADLER and ERIC A. POSNER**

ABSTRACT

Cost-benefit analysis is routinely used by government agencies in order to evaluate projects, but it remains controversial among academics. This paper argues that cost-benefit analysis is best understood as a welfarist decision procedure and that use of cost-benefit analysis is more likely to maximize overall well-being than is use of alternative decision procedures. The paper focuses on the problem of distorted preferences. A person's preferences are distorted when his or her satisfaction does not enhance that person's well-being. Preferences typically thought to be distorted in this sense include disinterested preferences, uninformed preferences, adaptive preferences, and objectively bad preferences; further, preferences may be a poor guide to maximizing aggregate well-being when wealth is unequally distributed. The paper describes conditions under which agencies should correct for distorted preferences, for example, by constructing informed or nonadaptive preferences, discounting objectively bad preferences, and treating people differentially on the basis of wealth.

COST-BENEFIT analysis (CBA) is widely and increasingly used by government agencies, yet academic debate about CBA is in disarray. Defenders of CBA traditionally conceptualize it as a technique for implementing either Pareto efficiency or Kaldor-Hicks efficiency, two criteria that many welfare economists take to be normatively basic. They also traditionally define Pareto efficiency, Kaldor-Hicks efficiency, and CBA in terms of a person's actual, as opposed to informed or otherwise undistorted, preferences. A project is (1) Pareto efficient relative to the status quo if at least one person actually prefers it to the status quo and no one prefers the status quo or (2) Kaldor-Hicks efficient relative to the status quo if there is a hypothetical

1105

costless redistribution from those who prefer the project to those who do not that would make the project Pareto efficient. A project is a CBA improvement over the status quo if the sum of compensating variations for the project is a positive number, where P's compensating variation (CV) for a project is the dollar amount paid to or from him in the project world such that, given P's actual preferences, he is indifferent between that world and the status quo. The academic critics of CBA argue that actual preferences are a poor basis for governmental policy. They further argue that the link between CBA and Pareto efficiency is tenuous and that the link between CBA and Kaldor-Hicks efficiency, although tighter, does not justify CBA, because Kaldor-Hicks itself lacks normative significance.

The critics are right, up to a point. The traditional defense of CBA is a failure. Where the critics go wrong is in thinking that this is a failure of CBA, rather than merely of a particular argument in its favor. As we have elsewhere tried to show, CBA is properly conceptualized as a welfarist decision procedure. The proper link is to the normative criterion of overall well-being, rather than to Pareto efficiency or Kaldor-Hicks efficiency, where overall well-being refers to the satisfaction of certain restricted preferences, rather than all actual preferences. (We will discuss the restrictions below.) By overall well-being we mean to include the large family of welfarist (but not necessarily utilitarian) moral and political theories, that is, those that hold that a policy's effect on people's welfare is a morally relevant, though not necessarily conclusive, consideration. The link between CBA and overall well-being is rough, not perfect: sometimes a project will be picked out by CBA as an improvement over the status quo and yet, in fact, be welfare inferior to the status quo. But CBA is sufficiently accurate in tracking overall well-being and has sufficient other procedural virtues— it is relatively cheap to implement, relatively easy to monitor by oversight bodies, and relatively undemanding of agency expertise—that it is plausibly the welfare-maximizing procedure for agencies to employ, in a significant fraction of their choice situations, compared to available alternative procedures.

The problem with the traditional definition of CBA in terms of actual preferences is that satisfaction of actual preference and maximization of well-being are not equivalent. But we think that the failure of the actual-preference view of well-being need not undermine CBA. In the course of defending this position, we will use the following vocabulary. Cost-benefit analysis can be redefined as the sum of welfare equivalents (WEs) rather than the sum of CVs, where P's WE is the amount of money such that— paid to or from him in the project world—he is just as well-off as in the status quo, on the correct theory of well-being. Thus defined, CBA is agnostic across theories of well-being and, in particular, is not committed to the

view that welfare and actual-preference satisfaction are equivalent. Because of this, when we discuss WEs, we will often not have a fixed number or formula in mind; the concept of WE (unlike the concepts of CV and equivalent variation, for example) is used as a placeholder to describe what dollar amount would be necessary under a given theory, or cluster of theories, of well-being. Nonetheless, even without making strong assumptions about the right theory of well-being, we will be able to derive some restrictions on how WEs are determined.

The concept of WE is necessary because the actual-preference theory of CBA must be abandoned. Actual preferences are not necessarily constitutive of welfare because they can be distorted, in various ways. For example, P may prefer the project to the status quo because he is uninformed; with fuller information, he would prefer the status quo. Or P's preference for the project may not track relevant criteria of objective value. The project harms innocent children; P, who is a sadist, prefers the project just because it does that, and he retains this despicable preference under full information. In these kinds of cases, traditional CBA equates the welfare impact of the project upon P to his CV, which is a positive number. Yet most people would agree that the ignorant or sadistic P, who prefers the project only in virtue of his ignorance or his sadism, should not be counted as a project beneficiary, who helps tip the cost-benefit scale in the project's favor. But how should the project's welfare impact upon P be calculated? At first glance, the correct agency response to the problem of distorted preferences seems to be to ask what P would hypothetically be willing to pay or accept if his preferences were relevantly undistorted (well-informed, nonsadistic) and use that number rather than P's CV in the cost-benefit calculus. The problem here is that P continues to have his actual preferences, not his hypothetical preferences, and his actual preferences—on the right theory of well-being—may have a significant effect on his welfare. If P remains ignorant and continues to prefer the project to the status quo, then the fact that he would (under full information) have a valuation of, say, $-\$100$ for the project hardly means that the actual welfare effect of the project is $-\$100$, and similarly for the sadistic P.

So agencies face a dilemma. On the one hand, CVs defined in terms of actual preferences deviate from welfare, when such preferences are distorted. On the other hand, hypothetical valuations in light of undistorted preferences may not capture the real welfare impact of projects either. This article takes a stab at solving that dilemma. The glib answer is that agencies should aggregate WEs, not CVs or hypothetical CVs. But this glib answer is really just a promissory note, because it remains to be seen how actual and undistorted preferences interact, for various kinds of distortions, to produce WEs. A further problem, which also needs to be confronted, is that

the sum of WEs can deviate substantially from overall well-being because the marginal utility of dollars depends upon the wealth of the person affected. (This is a well-known problem for conventional CBA based on CVs; the use of WEs does not by itself solve this problem.) The glib response here is that WEs should be weighted by a factor inversely proportional to wealth. As we discuss in more detail below, that response overlooks the point that rich and poor people behave in accordance with their actual wealth, not their hypothetical wealth.

Section I of the article lays out, in compressed form, our view of CBA as a welfarist decision procedure and further clarifies the idea of a WE.[1] Section II surveys agency practice and demonstrates that agencies performing cost-benefit analysis do, in fact, adjust valuations to correct for various kinds of preference distortions. Section III analyzes how agencies should adjust valuations to correct for preference distortions. We focus upon the five most important scenarios in which traditional CBA (the sum of unweighted CVs) and overall well-being may diverge: disinterestedness, poor information, adaptation to circumstances, objectively bad preferences, and wealth effects. For the first four scenarios, we discuss in detail how WEs can diverge from CVs and how WEs should be calculated; for the last scenario, wealth effects, we analyze various subtle problems associated with the form of CBA that weights WEs or CVs to reflect the declining marginal utility of income. Section IV addresses some institutional issues. Section V summarizes our views and presents concrete recommendations as to when agencies should depart from traditional CBA.

I. Cost-Benefit Analysis: The Traditional Defenses, and a New One

Cost-benefit analysis is traditionally linked to the familiar economic criteria of Pareto efficiency or Kaldor-Hicks efficiency. A project is Pareto efficient, relative to the status quo, if no one loses from the project and at least one person gains. A project is Kaldor-Hicks efficient, relative to the status quo, if there is a hypothetical costless lump-sum redistribution in the project world, from winners to losers, such that this amended project world is Pareto efficient relative to the status quo.[2] The concepts of welfare loss

---

[1] Section I is based on Matthew D. Adler & Eric A. Posner, Rethinking Cost-Benefit Analysis, 109 Yale L. J. 167 (1999), which contains further details and citations to the literature.

[2] Technically, there is a difference between Kaldor efficiency, Hicks efficiency, and Kaldor-Hicks efficiency. A project is (1) Kaldor efficient relative to the status quo if there is a hypothetical lump-sum redistribution in the project world, from project winners to project losers, such that this amended project world is Pareto efficient relative to the status quo; (2) Hicks efficient relative to the status quo if there is no hypothetical lump-sum redistribution in the status quo world, from project losers to project winners, such that this amended status quo world is Pareto efficient relative to the project; and (3) Kaldor-Hicks efficient if

and gain, as referenced by these criteria, are standardly cashed out in terms of preference satisfaction. P gains from a project, and loses from the status quo, if and only if he prefers the project. There are real difficulties in the equation of preference satisfaction and welfare, which we shall discuss in a moment. But even apart from such difficulties, the putative link between CBA and economic efficiency—either in the Pareto sense or in the Kaldor-Hicks sense—is one that needs to be broken.

Pareto efficiency has genuine normative import. An agency does the right thing, everything else equal, by approving a Pareto-efficient project. But the connection between CBA and Pareto efficiency is elusive. An individual project will be chosen by CBA over the status quo even though the project is not Pareto efficient relative to the status quo, as long as the aggregate CVs of those who gain from the project exceed the aggregate CVs of those who lose. The very idea behind CBA is to commensurate winners' gains and losers' losses, such that the project can be identified as better or worse even though neither option is Pareto efficient relative to the other. Nor can CBA be justified in light of Pareto efficiency by claiming (1) that project winnings will be redistributed through the tax system to project losers, since agencies employ CBA in many scenarios where no such redistribution will take place; (2) that a general policy of CBA is Pareto efficient for all affected, since agencies regularly employ CBA to impose losses sufficiently grave (for example, death or injury) that the persons thus affected would be better off in a world with no such policy;[3] and (3) that persons would choose a general policy of CBA if they were choosing ex ante and under conditions of uncertainty, since there are good arguments that persons in such circumstances would choose a different policy.[4]

As for Kaldor-Hicks efficiency, that criterion lacks genuine normative import. The fact that the project winners could compensate the project losers entails nothing of normative significance about the project. A project can be a Kaldor-Hicks improvement, relative to the status quo, without increasing overall well-being—for example, if the project produces a large

---

the project is Kaldor efficient and Hicks efficient. (The last criterion is also called the Scitovsky criterion.) For purposes of exposition, we adopt a simpler definition of Kaldor-Hicks efficiency in the text and focus upon what is, technically, Kaldor efficiency. Our arguments readily carry over to Hicks efficiency and to Kaldor-Hicks efficiency in the technical sense; none of the three criteria has genuine normative import.

[3] Note further that, because CBA inflates the welfare effect of projects on rich persons and deflates the welfare effect on poor persons, poor persons might also be better off in a world without the practice of CBA.

[4] John Harsanyi famously argued that persons choosing ex ante and under conditions of uncertainty would choose a policy of welfare maximization. See John C. Harsanyi, Morality and the Theory of Rational Behavior, in Utilitarianism and Beyond 39 (Amartya Sen & Bernard Williams eds. 1982). But welfare maximization is different from CBA.

increase I in the resource endowment of richer persons and a smaller decrease D in the resource endowment of poorer persons, such that $I > D$ (making the project Kaldor-Hicks efficient), but the welfare impact of I is less than the welfare impact of D. Nor should Kaldor-Hicks be thought to have normative significance apart from welfare maximization. For example, any normative link to the powerful notion of consent is vitiated by the crucial point that it is hypothetical winner-to-loser compensation, not actual winner-to-loser compensation, that makes a project Kaldor-Hicks efficient. The fact that the losers would or should consent to the efficient project, if the redistribution from winners to losers were performed, does not mean that the losers do or should consent to the project where no such redistribution actually takes place.

In short, the fact that a project with a positive total sum of CVs, relative to the status quo, is Kaldor-Hicks efficient relative to the status quo does not explain why the agency ought to choose the project. Further, it turns out that—for technical reasons—a project can have a positive, total sum of CVs relative to the status quo and not be Kaldor-Hicks efficient.[5]

Cost-benefit analysis needs to be reconceptualized. The way to do so is twofold. First, CBA itself must be recognized to lack normative significance.[6] The fact that a project has a positive sum of CVs does not mean that it is a genuine moral improvement over the status quo.[7] Rather, CBA is a decision procedure. It is a technique used by agencies for choosing between options, a technique whose justifiability must be evaluated in light of normative criteria with which CBA is only contingently connected.

Second, there is a genuine normative criterion that does plausibly justify the use of CBA, and that is the criterion of overall well-being. It is this criterion, not Pareto efficiency or Kaldor-Hicks efficiency, that provides the normative foundations for CBA. Overall well-being is one of the several normative criteria bearing upon governmental choice; it is morally relevant

[5] See Robin W. Boadway, The Welfare Foundations of Cost-Benefit Analysis, 84 Econ. J. 926 (1974).

[6] Perhaps it is an exaggeration to describe this particular element of our view of CBA as a "reconceptualization." Sophisticated contemporary defenders of CBA, within welfare economics, view it as a decision procedure that imperfectly implements either Pareto efficiency or Kaldor-Hicks efficiency. But our overall view of CBA clearly is a reconceptualization; and seeing CBA as a decision procedure rather than a bedrock normative criterion is an important component of our overall view.

[7] The reason that CBA lacks normative significance is parallel to the reason that Kaldor-Hicks efficiency does. A project may produce gains for the rich (whose CVs are inflated by their wealth) and losses for the poor (whose CVs are deflated by their poverty), such that it is counted as an improvement by CBA even though the project decreases overall well-being. So CBA is not equivalent to the welfare criterion. And, as with Kaldor-Hicks efficiency, it is very hard to see how CBA could have moral significance by tracking or constituting some moral criterion other than the welfare criterion.

to, if not conclusive of, government's choice between a project and the status quo. Thus we disagree with classical utilitarianism, which is the view that overall well-being is the sole criterion bearing on governmental choice. Our view is that the government should choose a welfare-improving project unless other considerations, such as deontological or egalitarian considerations, justify rejecting the welfare-improving project and choosing the status quo instead. How such nonwelfarist considerations should be operationalized at the agency level is a complicated matter. It may be the case that an institutional division of labor obtains, where agencies focus solely on the welfare criterion and courts and legislatures then revise agency choices in light of deontological criteria, egalitarian criteria, and the like. In any event, we do not mean to suggest that CBA is a superprocedure, which implements all the moral considerations bearing upon governmental choice. Rather, CBA is plausibly justified in light of a particular moral criterion— overall well-being—and thus is plausibly one part of the set of procedures and institutions comprising government.

Note that our view, albeit nonutilitarian, does entail a commitment to the possibility of interpersonal welfare comparisons. It was once a widely held view within welfare economics that such comparisons are impossible—that Pareto-noncomparable states are welfare noncomparable. That view has changed and, in any event, is wrong. Consider a project that causes minor headaches to a few people, but averts many premature deaths. The project is Pareto noncomparable with the status quo, but surely it increases overall well-being. In our ordinary lives, as family members or citizens, we routinely judge that the positive (or negative) welfare effect of an option on some person or persons is large enough to outweigh the negative (or positive) welfare effect of the option on others. Any theory of well-being that does not license such comparisons is, on those very grounds, an unreasonable theory.[8]

Granting the moral relevance of overall well-being and the possibility of interpersonal welfare comparisons, why think that CBA is a decision procedure justified in light of overall well-being? There are a variety of possible decision procedures by which agencies might choose between their options. One such procedure is CBA; another procedure is the direct implementation

---

[8] Skeptics of interpersonal welfare comparisons sometimes point out that it is not practicable for government to make such comparisons. We agree. But this is not the same as saying that interpersonal comparisons are impossible. To think otherwise is to conflate decision procedures and normative criteria. Although actual agencies should not choose between project and status quo by attempting to perform a comparison of the winners' welfare gains with the losers's losses—because that procedure is too expensive, and so on—there remains, or may remain, a right answer to the question, Are the winners' welfare gains larger than the losers' losses?

of the welfare criterion, where agencies actually attempt to judge whether overall well-being is greater in the project world or the status quo; yet another procedure is one where agencies pick a single dimension bearing upon welfare (such as longevity, or aesthetic experience, or employment) and choose options by maximizing along that single dimension. Or an agency might eschew maximization and aggregation entirely, as agencies do when they make choices by looking to social norms. Why think that CBA, by contrast with these other procedures, is the best procedure for agencies to use—the procedure the use of which maximizes overall well-being? Our view is actually a bit more qualified: although there are some scenarios in which agencies should not employ CBA (for example, where the project's likely impact on overall well-being is too small to warrant the expense of CBA), CBA is welfare justified in a significant fraction of agency choice situations. By contrast with unidimensional and nonaggregative procedures, CBA is relatively accurate—although not perfectly accurate—in tracking overall well-being. By contrast with direct implementation of the welfare criterion, CBA is cheaper to perform, less prone to agency error, and more readily monitored by oversight bodies such as legislatures or the Office of Management and Budget (OMB).

These are surely contestable and (to some extent) empirical claims, which bear further examination and debate. But at a minimum it is clear to us what the debate about CBA should concern: it should concern whether, as compared with other possible agency procedures, CBA is the procedure by which the welfare criterion is best implemented. Again, it is welfare, not efficiency, that figures here, and it must be kept firmly in mind that although CBA itself lacks normative import, its use as a decision procedure may be better justified in light of some genuine normative criterion than the direct implementation of that criterion.

The reader, by this point, may have noticed one large lacuna in our sketch of the welfarist justification for CBA. That concerns the link between CBA and actual preferences. Cost-benefit analysis is traditionally defined as the sum of CVs, and CVs in turn are defined in terms of actual preference. P's CV is the dollar amount such that, paid to or from him in the project world, he is indifferent—given his actual preferences—between that world and the status quo. But there are many reasons to think that an actual-preference account of welfare is wrong. At a minimum, that account is wrong because it treats disinterested or morally motivated preferences as welfare relevant. A preference is simply a particular kind of ranking (specifically, a ranking by P that has some explanatory connection to P's choices). But if P prefers (ranks) the project over the status quo simply because he believes the project to be morally required, then surely the occurrence of the project need not benefit P. To think otherwise is to confuse the

totality of factors that influence P's rankings and choices—including his moral views, his special commitments to family and friends, and so forth—with one such factor, namely, P's welfare.[9]

An example might help make the point. Imagine that the project, which would be implemented in Montana, involves eliminating an endangered species located in that state. P, who lives in New York, believes that this is morally wrong because he believes that endangered species have intrinsic value. P has no other involvement with the species or Montana. He has never been to that state and never plans to go. His professional work has nothing to do with the environmental movement, and he has no particular nonprofessional involvement either: he does not devote large amounts of his time, or resources, or attention to environmental issues. Still, he has been persuaded (after reading a book on environmental ethics) that there are strong moral grounds against eliminating endangered species. He therefore rejects the project; he ranks it below the status quo, and this ranking is connected to his choices, in the sense that it leads him (let us imagine) to take certain steps to help the species, such as sending a $100 check to the Sierra Club. This hardly entails that P personally benefits from the status quo—that the project harms his own welfare. After all, P himself would deny that claim, and it would run afoul of our commonsense intuitions about welfare. P rejects the project not because he believes it harms him but because he believes that the project is morally wrong; and there is nothing driving P's preference, no personal or professional connection, other than this belief. Thus P's act of sending the check to the Sierra Club is disinterested: it is an action that diminishes P's own well-being (because he loses $100) and that he takes to do so, but that is motivated by considerations other than P's own well-being. If welfare and preference satisfaction are equivalent, then the concept of disinterested action is incoherent.

We therefore adopt a restricted preference-based account of well-being. A project is a welfare improvement for P only if P prefers the project and only if this preference is properly restricted to P's own interests, concerns, and welfare. Providing a precise account of this restriction is difficult, but it is clear that some such amendment to the actual-preference account of welfare is needed—as the case of morally motivated preferences shows.

More generally, there are a variety of factors, other than actual preference, that may bear upon P's welfare. These factors include the following:

---

[9] Sen distinguishes preferences based on "commitment"; Harsanyi distinguishes "ethical preferences." See Amartya K. Sen, Rational Fools, 6 Phil. & Pub. Aff. 317 (1977); John G. Harsanyi, Cardinal Welfare, Individualistic Ethics, and Interpersonal Comparisons of Utility, 63 J. Pol. Econ. 315 (1955). Other works defending the view that only a restricted set of preferences are relevant for well-being—that some preferences can be disinterested—are cited in Adler & Posner, *supra* note 1, at 197–204.

Information.   P prefers the project. However, with fuller information, P would prefer the status quo. Arguably, P is not benefited by the project, or at least not benefited as much as he would be if his fully informed preferences were in favor of the project.

Objective Good.   P prefers the project. However, the status quo is objectively much better for P, in light of objective welfare goods such as friendship, knowledge, aesthetic experience, or accomplishment. Arguably, P is not benefited by the project, or at least not benefited as much as he would be if the objective value of the project were greater.

Adaptive Preference.   P prefers the status quo because that is the world in which his preferences were formed. However, P's preferences are misshapen by various unjust features of the status quo. (Imagine that the status quo is a world in which P lacks wealth, or self-respect, or a basic education, and in which his preferences have been shaped by these deficits.) If the project is implemented, P's preferences may change—such that he now prefers the project, not the status quo—or the preferences may be sufficiently entrenched that they do not change. In either event, there may be good grounds for thinking that P benefits from (or at least is not harmed by) the project even though he prefers or once preferred the status quo.

Affect and Experience.   P prefers the project. However, if the project were implemented, he would not enjoy it. Or, if the project were implemented, he would not experience—become aware of—that fact. (Note that P's preference for a state of the world is satisfied if the state of the world occurs; it is a further and contingent matter whether P experiences the occurrence of the state of the world.) Arguably, P cannot be benefited by a project that he does not enjoy or experience.

We are not sure that the actual-preference view of well-being needs to be modified in all the ways just listed. These are plausible proposals, but we find them somewhat less compelling than the earlier point that morally motivated and other unrestricted preferences are welfare irrelevant. On the other hand, we are hardly confident that the right theory of well-being omits reference to P's information, to the objective goods that he realizes, to the adaptive cast of his preferences, or to his affect and experience. Thus we need to provide a defense of CBA that is consistent with such amendments to an actual-preference view of welfare.

One defense returns to the point that CBA is a decision procedure, not a normative criterion. P's CV does not perfectly capture the welfare effect of the project upon him, given the disjunction between the actual-preference view of welfare and the correct view—a restricted-preference view that may further incorporate factors such as information, objective goods, adaption, and affect or experience. But perfect capture is not needed. Compensating variations are sufficiently accurate in tracking welfare impacts and

sufficiently easy to implement and monitor by agencies that CBA as traditionally defined is welfare justified in light of the right theory of welfare, notwithstanding the disjunction from the actual-preference view. Or so the argument in favor of traditional CBA might go.

This argument may be persuasive, at least in part. There may well be scenarios in which traditionally defined CBA is indeed the welfare-maximizing decision procedure for agencies to employ. But it is important to see that the welfarist defense of CBA is not limited to the argument just sketched. Cost-benefit analysis need not be defined in the traditional way, as the sum of CVs. The valuation concept at the core of CBA can be amended. As we noted earlier, CVs might be replaced with welfare equivalents, or WEs, where P's WE is the amount of money paid to or by him in the project world such that—on the right theory of well-being—P is just as well-off there as in the status quo. Note that CBA, defined as the sum of WEs, automatically corrects for the disjunction between the actual-preference view and the right theory, whatever that theory happens to be. For example, imagine that the view is a hybrid view, such that P benefits from the project if and only if (1) he actually prefers it and (2) he would prefer it under full information. In the case where ignorant P prefers the project but fully informed P would prefer the status quo, P's CV is a positive number but—on the theory at hand—the welfare impact of the project upon him is nil. And so P's WE is also nil, because that concept (by contrast with the simpler concept of the CV) makes reference to the full and correct welfare theory.

To be sure, the accuracy of the WE in tracking the right welfare theory is, from another point of view, a defect. Compensating variations are relatively easy to calculate; WEs may not be. Telling agencies to aggregate WEs rather than CVs may give them additional scope for shirking and error. But we are confident that, at least in some cases, CBA defined as the sum of WEs (or something like that) will be welfare maximizing, as compared to traditional CBA and other procedures. Cost-benefit analysis can, to some extent, be modified in a way that corrects for the failings of an actual-preference view of welfare and that still leaves in place a practicable choice procedure. At a minimum—since the traditional version of CBA cannot just be assumed to be the welfare-maximizing version—economists, lawyers, and philosophers need to begin considering what a sum-of-WEs approach would involve. That is the task to which we turn in the remainder of this article.

One final and significant point bears mention here. We have criticized an actual-preference theory of welfare for making P's preferences the sole constituents of his welfare. The right theory adds additional elements—restrictions, information, objective goods, to name some plausible candi-

dates—beyond sheer preference. But preference cannot be dispensed with entirely. The right theory of well-being, in our view, is one that gives preference[10] a partial and constitutive role; theories that fail to do so, so-called objective or hedonic theories, are therefore mistaken. If P prefers the status quo to the project, and would do so (and continue to do so) even if the project were implemented, then P is not benefited by the project. The fact that P hypothetically would approve the project under full information, or that it is objectively good, may be necessary conditions for the project's benefiting P, but they are not sufficient conditions. To think otherwise is to give too little weight to P's own point of view; it is to think, implausibly, that P can made better off by an option in the teeth of P's actual and continuing aversion to it. This will be of much importance in thinking about how agencies should modify CBA.[11]

II. HOW GOVERNMENT AGENCIES CORRECT
FOR DISTORTED PREFERENCES

A preference distortion arises when a person's CV for a project does not accurately measure the extent to which the project improves his welfare, properly understood. The previous section argued that preference distortions may have diverse causes. A person's CV for a project is inaccurate if it reflects his disinterested preferences rather than his self-interested preferences. Further, a person's CV for a project may be inaccurate if his preferences are uninformed, adaptive, or objectively bad. Finally, a person's CV for a project is inaccurate if it is inflated by his relative wealth or deflated by his relative poverty.

The preference distortions identified in the prior section may seem more theoretical than real. However, in this section we argue that government agencies act in ways that reflect concerns about preference distortions, and they therefore deviate from what we call "textbook CBA"[12] (by which we

[10] More precisely, the right theory of well-being is one that gives preference or some other pro-attitude a partial and constitutive role. Thus in our earlier paper, see Adler & Posner, *supra* note 1, at 197–204, we characterize ourselves as adopting a restricted desire-based view, where desires are, generically, pro-attitudes, including, specifically, preferences. This is a point of detail that need not be pursued here.

[11] More precisely, P can only be intrinsically benefited by a project if he prefers it or comes to prefer it. Clearly, P can be instrumentally benefited absent a conforming preference, for example, in the case where P prefers a pill that he believes to be a health-causing vitamin, when in fact it is a death-causing poison. The "project" of keeping the pill from P benefits him notwithstanding his unchanging belief that the pill is a vitamin. See Adler & Posner, *supra* note 1, at 202 n.97.

[12] We acknowledge that CBA textbook authors sometimes discuss these distortions. By "textbook CBA," we mean the kind of CBA that is generally espoused by its advocates, that is, the sum of unweighted CVs based upon actual preferences. Citations to the economic literature on CBA, and in particular to the literature advocating CBA, are provided by Adler & Posner, *supra* note 1, at 187–94.

mean, again, the sum of unweighted CVs based on actual, unrestricted preferences).

In Section III below, we will evaluate these administrative actions and propose modifications of the ways in which agencies deal with preference distortions. We should stress at the outset that we do not approve of all, or even many, of the government's approaches to these problems; we spend time on them in order to show that these problems are real and pose important practical difficulties.

A. Disinterested Preferences: The Problem of Existence Value

Textbook CBA reduces moral commitments to valuations. Consider, for example, the recent debate over the use of contingent valuation methods to value environmental goods. Textbook CBA, as generally understood, directs agencies to translate people's moral attitudes about the environment into CVs for the existence of environmental goods that they do not directly enjoy, usually called "existence value" or "nonuse value."[13] These CVs are then added to the balance of costs and benefits of a project, like any other CV.

Until recently, agencies did not calculate existence values and use them in order to evaluate regulations. The earliest sustained discussion of existence values by an agency that we have found occurred in 1986 and involved the Department of Interior's guidelines on valuing environmental damage caused by a discharge of oil or a hazardous substance. But the rule itself did not involve the calculation of an existence value.[14] The first use of existence values in rule making was, as far as we have found, by the Environmental Protection Agency (EPA) in 1991,[15] and the practice of measuring existence values has only recently become common. For example, the EPA's recent CBA for effluent regulations included existence valuations for "benefits to wildlife, threatened or endangered species, and biodiversity benefits."[16] Thus, the use of existence values by government

---

[13] Existence value is the value from knowing that some good exists; it is sometimes used interchangeably with nonuse value, but nonuse value also is understood to include the option value of having some good in the future, which we exclude from our analysis.

[14] See Department of the Interior, Natural Resource Damage Assessments, 51 Fed. Reg. 27,674 (1986).

[15] See Environmental Protection Agency, Approval and Promulgation of Implementation Plans: Revision of the Visibility FIP for Arizona, 56 Fed. Reg. 5173 (1991).

[16] Environmental Protection Agency, Effluent Limitations Guidelines, Pretreatment Standards, and New Source Performance Standards for the Transportation Equipment Cleaning Point Source Category, 63 Fed. Reg. 34,686, 34,724 (1998). See also Environmental Protection Agency, Lead Fishing Sinkers: Response to Citizens' Petition and Proposed Ban, 59 Fed. Reg. 11,122, 11,135 (1994) (endorsing use of existence values). The use of the methodology has been approved by the D.C. Circuit, see Ohio v. U.S. Dep't Interior, 880 F.2d 432, 474–81 (D.C. Cir. 1989).

agencies has lagged the widespread use of cost-benefit analysis by about a decade.

One reason for hesitation about calculating existence values was no doubt methodological. Existence values cannot be inferred from market behavior, but must be derived from costly and controversial surveys. Another reason for hesitation might have been politics. But it is also likely that agencies have been uncertain about the conceptual soundness of using existence values. If not, they should have been—as we argue in the next section.

Whether or not we are correct about the reasons for the delay in using existence values in environmental regulation—namely, that existence values do not constitute morally relevant information, either with respect to overall well-being or with respect to other moral criteria—we think these reasons do explain why existence values are not used outside environmental regulation. The Food and Drug Administration (FDA) does not ask Christian Scientists whether they care about the existence of people using commercial drugs. The U.S. Department of Agriculture (USDA) does not ask animal rights activists whether they care about the existence of slaughterhouses. The U.S. Postal Service does not ask individuals whether they care about the existence of pornography in the mail. One might be able to point to some factors that distinguish these cases from environmental regulation. The Constitution bars the government from implementing religious views. Concerns about the treatment of farm animals are not as widespread or deep as concerns about the environment. But there is no conceptual reason for distinguishing among these different contexts. If people's disinterested preferences are worthy of consideration in CBAs, then disinterested environmental preferences are not the only ones that matter. All such attitudes—including attitudes toward abortion, the death penalty, medical research, family structure, the treatment of children, and the appropriateness of government intervention—should be quantified, monetized, and weighed against opposing costs and benefits when agencies implement projects.

B. Uninformed Preferences

Either textbook CBA ignores the problem of uninformed preferences (by assuming that all persons have perfect information), or it recognizes the problem but purports to solve it by conceptualizing information as yet another good. Individuals know that they lack information and are willing to pay for more information if and only if the expected gain exceeds the cost. So an individual who is partly uninformed about a given project is rationally uninformed, and his CV for the project should be calculated based upon his uninformed preferences.[17]

[17] In particular, the individual's CV would be set equal to the expected value of the project or to the single ex post payment in all the project states of the world such that (given the

Government agencies' treatment of uninformed preferences is, in fact, more complex. At the outset, it is worth distinguishing between the effect of information on (i) instrumental preferences and on (ii) intrinsic preferences. In the first case, information has the effect of changing persons' judgments about the causal link between the project and those states of affairs that they intrinsically value or disvalue. P opposes a project to fluoridate drinking water because he falsely believes that fluoridation causes cancer and does not help teeth. In the second case, information changes intrinsic valuations. P's CV for an arts project is low because he has not been fully informed about the aesthetic qualities of the project.

Agencies frequently refuse to use CVs that reflect uninformed preferences. Sometimes, agencies supply people with information when asking them for their CVs. Before asking them about air quality over the Grand Canyon, the EPA showed survey respondents photographs of the site with different levels of pollution.[18] The EPA's goal was presumably to provide information on environmental aesthetics, about which respondent's intrinsic preferences were uninformed. For a regulation involving labeling of meat and poultry products, the USDA relied on CVs for health benefits people would enjoy if they altered their behavior in response to the labels, rather than people's CVs for nutrition disclosure.[19] The agency appeared to take the intrinsic preference (for health) as a given and to circumvent the problem of imperfectly informed instrumental preference (for nutritional disclosure). Instrumental preferences are constructed; people are assumed to have preferences for whatever means will best satisfy their intrinsic preferences, even if they are misinformed about these means—even if, in our example, some people would oppose labeling because they (falsely) think it will confuse them.

The premise of modern workplace regulation is that workers are uninformed about risks. If this premise is false, then wages and workplace safety procedures reflect rational trade-offs made by workers, and regulations simply interfere with the satisfaction of their preferences. It would make no sense, for example, for the Occupational Safety and Health Administration to restrict workers' exposure to ethylene dibromide, a carcinogenic chemical used in various industries,[20] because any restrictions suffi-

---

individual's information) he is indifferent between the project and the status quo. Lewis Kornhauser's contribution to this symposium provides a more detailed discussion of how textbook CBA calculates CVs when actors are not fully informed. See Lewis A. Kornhauser, On Justifying Cost-Benefit Analysis, in this issue, at 1037.

[18] See Leland B. Deck, Visibility at the Grand Canyon and the Navajo Generating Station, in Economic Analysis at EPA 267 (Richard D. Morgenstern ed. 1997).

[19] Department of Agriculture, Nutrition Labeling of Meat and Poultry Products, 56 Fed. Reg. 60,302 (1991).

[20] See Department of Labor, Occupational Exposure to Ethylene Dibromide, 48 Fed. Reg. 49,959 (1983).

ciently cheap that workers are willing to pay for them have voluntarily been implemented by employers. More risk-averse workers would move to safer, lower paying jobs; more risk-preferring workers would take the more dangerous, higher paying jobs. In such a labor market, regulations would be costlier than their safety benefits warrant. The denial that this market prevails is implicit in all regulation of contractual relations; agencies occasionally are explicit about it.[21]

In sum, agencies do not always take uninformed preferences as they find them. Instead, they sometimes evaluate projects by using the preferences people would have if they were informed.

C. Adaptive Preferences

Textbook CBA does not recognize the existence of adaptive preferences. Preferences that are the result of adaptation are treated the same as preferences that are not the result of adaptation.

As noted in Section I, people may psychologically adapt to an unjust or otherwise unfavorable environment, so that their CV for eliminating a risk or irritant is less than what it would be if they did not adapt. An overburdened housewife might rationalize her position and so not be willing to pay in order to have her burdens removed.[22] Or a person in a bad environment might feel sour grapes toward someone in a pleasant environment and refuse to pay for an improvement in his own environment because he has convinced himself that the pleasant environment is really worse. Or a person might adapt to the status quo, so while he will oppose any project, if the project were nevertheless implemented, he would oppose a further project that would reverse the first.[23]

Many agency programs assume that people's preferences are distorted by psychological problems. For example, programs to reduce drug use assume that drug addicts would be benefited by restrictions on drugs, not harmed by them, even though their preferences may well be the opposite. In justifying regulations governing antidrug programs for the employees of private air carriers, the Department of Transportation (DOT) did not take into account the preferences of the drug users, even though these people may well be hurt by the regulations on an actual-preference account of CBA.[24] In justi-

---

[21] See, for example, Department of Labor, Occupational Exposure to Bloodborne Pathogens, 56 Fed. Reg. 64,004, 64,087 (1991) (arguing that workers do not know about many risks and are unable to analyze them correctly).

[22] Amartya Sen, On Ethics and Economics 45 (1987).

[23] Cass R. Sunstein, Free Markets and Social Justice 252–53, 256–58 (1997); Jon Elster, Sour Grapes: Studies in the Subversion of Rationality (1983).

[24] Department of Transportation, Anti-drug Program for Personnel Engaged in Specified Aviation Activities, 53 Fed. Reg. 47,024 (1988).

fying mandatory drug tests for drivers of commercial vehicles, the DOT did not take account of the cost to drivers who derive pleasure from the use of alcohol and illegal drugs.[25] Yet in these same regulations, the agencies did take account of preferences that are not considered adaptive, such as preferences for time or money.

However, several qualifications are necessary. First, agencies do not generally assume that preferences are defective in these ways unless directed by a statute. Second, it is not clear whether the preferences of drug users are ignored because they are adaptive or because they are considered objectively bad (see below) or distorted in other ways.[26]

D. Objectively Bad Preferences

Textbook CBA assumes that objectively bad preferences should receive the same weight as morally neutral preferences. An agency should presumably count the preferences of a person who hates children, and is willing to pay $1,000 to prevent a children's vaccine program, or the preferences of a person who hates homosexuals, and is willing to pay $1,000 to prevent AIDS research.

Agencies routinely ignore sadistic preferences and other objectively bad preferences. For example, the FDA's cost-benefit analysis of a regulation designed to curb distribution of cigarettes to children did not include as a cost the lost profits to industry, "because most of this profit stems from illegal sales to youths."[27] Nor did it count the children's lost consumer surplus. It is hard to believe that agencies would count a preference that homosexuals not be helped through AIDS research, no matter how widespread that preference may be. And, as we saw above, the DOT's refusal to count the preferences of drug users may reflect an evaluative judgment (on the

[25] See Department of Transportation, Federal Motor Carrier Regulations; Controlled Substances and Alcohol Use and Testing; Commercial Driver's License Standards, Requirements and Penalties; Hours of Service of Drivers, 57 Fed. Reg. 59,567 (1992).

[26] Rather than being adaptive or objectively bad, it might be the case that a drug user's preference is (1) uninformed, in the sense that she is mistaken about the consequences of drug use (health consequences, the risk of addiction); (2) akratic, in the sense that the drug user prefers not to use drugs, given their consequences (of which she is aware), but nonetheless irrationally continues to use them; (3) conflicted, in the sense that she has a first-order preference for drug use but a second-order preference not to have this first-order preference; or (4) coerced, in the sense that the drug addict has a preference for drug use only because she will experience miserable withdrawal symptoms if she stops using drugs. Compare Douglas N. Husak, Drugs and Rights (1992) (considering, but rejecting, possible grounds for regulating recreational drug use).

[27] Food and Drug Administration, Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents, 61 Fed. Reg. 44,396, 44,593 (1996).

part of Congress) rather than, or in addition to, a concern that the preferences are adaptive.

As another example, consider a Federal Aviation Administration (FAA) program for airline security, which refused to use profiling on the basis of race, national or ethnic origin, and other possibly relevant but morally suspect factors. Although the FAA performed a CBA, it did not consider the possibility that a discriminatory system might be less costly than the system that it endorsed[28] or that some people (white supremacists, racially biased airline travelers) might have strong tastes for discrimination and therefore have significant, positive CVs for a discriminatory system.

In sum, agencies depart from textbook CBA by refusing to weigh certain kinds of objectively bad preferences. This practice is not as obvious as other adjustments are, because we are not accustomed to thinking that satisfying preferences for discrimination, suffering, and other morally bad outcomes will benefit the holders of the preferences. Thus the agencies' practice seems natural. But that is only because in this significant respect textbook CBA deviates from common moral intuitions.

E.   Wealth Distortions

Textbook CBA does not adjust for distortions caused by the distribution of wealth. A wealthy person is willing to pay more to reduce the risk of death than is a poor person with identical preferences, but it does not follow that the agency maximizes welfare by placing dangerous projects in poor neighborhoods rather than in rich neighborhoods. On the contrary, it seems reasonable to assume that premature death has the same effect upon overall well-being, whether the person who dies prematurely is wealthy or poor. Note that this distortion is different from the other three. The others arise because of the disjunction between preference satisfaction and well-being; relatedly, they cause (or may cause) a deviation between CVs and WEs. By contrast, the wealth distortion described just now arises because CVs and WEs are calculated in terms of dollars, which do not accurately reflect relative well-being when endowments differ. This distortion would exist even if actual and undistorted preferences did not diverge, and (where they do diverge) even if agencies were able to costlessly and accurately calculate WEs.

Agencies correct for wealth distortions in various ways. They use a constant figure for the monetized value of life.[29] They rely on quality-adjusted

---

[28] Department of Transportation, Security of Checked Baggage on Flights within the United States, 64 Fed. Reg. 19,220 (1999).

[29] See U.S. General Accounting Office, Regulatory Reform: Agencies Could Improve Development, Documentation, and Clarity of Regulatory Analysis 26–27 (GAO/RCED-98-142,

or -nonadjusted life-years, which is a number that is invariant with wealth.[30] The Department of Health and Human Services, for example, said that the benefit of its organ transplant rule was the saving of 297–1,306 life-years.[31] They quantify other benefits without monetizing them. The EPA, for example, noted that a regulation of certain heavy-duty engines would reduce nitrogen oxide emissions by 593,000 tons, without attaching a value to this amount.[32] In all these cases, the benefits of the regulation are invariant to the wealth of those affected by them. Of course, as the agencies depart farther and farther from the use of CVs, the basis of evaluation becomes increasingly obscure.

Agencies also correct for wealth distortions in more broad-gauged ways. Regulation of pesticides and lead-based paint may have been influenced by a desire to benefit, on distributional grounds, low-income farm workers, in the first case, and low-income inner-city residents, in the second.[33] Agencies also pay attention to whether a regulation will "threaten the existence" of an industry, that is, have substantial, concentrated impacts.[34] Indeed, Clinton-era executive orders that require consideration of "environmental justice" and equity appear to require attention to the distributional consequences of regulations.[35] The original executive order directing agencies to use CBA says: "When an agency determines that a regulation is the best available method of achieving the regulatory objective, it shall design its regulations in the most cost-effective manner to achieve the regulatory objective. In doing so, each agency shall consider incentives for innovation, consistency, predictability, the costs of enforcement and compliance (to the

---

1998). More precisely, they use a range, but they do not make their choice within the range depend on the wealth of the victims.

[30] See Department of Health and Human Services, Regulatory Impact Analysis of the Proposed Rules to Amend the Food Labeling Regulations, 56 Fed. Reg. 60,856, 60,871 (1991) ("dying of a heart attack at age 80 is posited to be of less societal concern than dying in a car accident at age 35"). Yet the 80-year-old might have a higher CV than the 35-year-old.

[31] Office of Management and Budget, Draft Report to Congress on the Costs and Benefits of Federal Regulations, 63 Fed. Reg. 44,034, 44,048 (1998). See also id. at 44,052–53, on the FDA and other agencies.

[32] Id. at 44,048.

[33] Louis P. True, Jr., Agricultural Pesticides and Worker Protection, in Economic Analysis at EPA 324 (Richard D. Morgenstern ed. 1997); Environmental Protection Agency, Lead: Identification of Dangerous Levels of Lead, 63 Fed. Reg. 30,302, 30,305 (1998); Department of Housing and Urban Development, Office of Lead-Based Paint Abatement and Poisoning Prevention; Requirements for Notification, Evaluation and Reduction of Lead-Based Paint Hazards in Federally Owned Residential Property and Housing Receiving Federal Assistance, 61 Fed. Reg. 29,170, 29,202 (1996).

[34] See, for example, Department of Labor, supra note 21, at 64,082.

[35] See Exec. Order No. 12,898, 59 Fed. Reg. 7629 (1994) (environmental justice).

government, regulated entities, and the public), flexibility, *distributive impacts, and equity*."[36]

One might doubt whether these instructions have had much impact. On the one hand, agencies typically publish, alongside the CBA, a discussion of distributive impacts. On the other hand, these discussions are usually formulaic and inconclusive. But they do show that distributive issues have some prominence, contrary to the prescription of textbook CBA.[37]

F. Some Objections

Agency practice diverges in many ways from the requirements of textbook CBA. We argue that these practices reflect, in a loose and mostly unarticulated way, our concerns about the moral foundations of CBA.

An alternative hypothesis is that the divergence between agency practices and textbook models reflects either attempts to economize on decision costs or political constraints.

Neither of these alternatives is plausible. Agencies economize on decision costs in many ways, but what needs to be explained is a more complex pattern of behavior. Often agencies do expend considerable resources to determine what people's CVs are, and often they do not; the question is, why do they expend resources sometimes and not at other times? Consider agencies' refusals to determine the CVs of drug users when a project increases the cost of using drugs or their refusals to determine the CVs of racists for a project that involves racial profiling. Calculating these CVs is no more costly than calculating CVs in other contexts—for example, the CV of someone who is harmed by airplane noise or the CV of someone who experiences an increased risk of cancer. When agencies face high decision costs, they are usually quite candid about this problem. They will say that data are unavailable or too costly to acquire. They will make estimates based on a small sample. It is impossible to imagine an agency saying that the only reason that it did not calculate the CVs for drug use or racial discrimination is that data were unavailable.

The political argument might be that interest group politics result in the divergences that we have discussed. But there are interest groups on both

---

[36] Exec. Order No. 12,866, 58 Fed. Reg. 51,735 (1993) (emphasis added). Guidelines issued by OMB say, "When benefits and costs have significant distributive effects, these effects should be analyzed and discussed, along with the analysis of net present value." Office of Management and Budget, Guidelines and Discount Rates for Benefit-Cost Analysis of Federal Programs, Circ. No. A-94, Rev. (October 29, 1992), available from Westlaw in the database "OMB-circular."

[37] Ironically, the one thing that agencies do not do is use explicit weightings of CVs on the basis of marginal utilities of money, which is the most popular approach in CBA textbooks.

sides of virtually every regulation. A more plausible political account of agency practice is that agencies fear public outrage. Agencies hesitate about using wealth-dependent valuations of life, or respecting preferences for drug use, or attaching dollar values to environmental amenities, because these practices may produce public outrage. Public outrage may reflect ignorance, or herding, or strategic behavior,[38] but we think that it also reflects a conviction that deep moral intuitions are being ignored. If this is the case, the political argument is no different from our moral argument.

Defenders of textbook CBA will more likely regret the divergence between textbook CBA and agency use of CBA than explain it away. But we think that the divergence justifies attention to the problems with textbook CBA and provides support for our view that unrestricted preferences should not always be the basis of agency action. Nonetheless, the agency practices we have discussed raise questions.

First, as we have already pointed out, it is improper for agencies to correct for the fact that actual preferences may be distorted by shifting wholesale to undistorted preferences and calculating CVs based on those. What kind of shift should be made is a subtle problem and will vary depending on the source of distortion. Second, the modifications risk making CBA less transparent, cheap, and reliable. When agencies are told that the preferences relevant for CBA are different from the preferences people actually hold, they obtain a degree of freedom that will interfere with review by hierarchical superiors. Finally, there is no generally accepted means for correcting for wealth distortions. The balance of the article examines these problems in detail.

III. How Agencies Should Respond to Distorted Preferences

A. Disinterested Preferences

Many people criticize agencies' use of contingent valuation techniques to measure existence values, but these criticisms are for the most part methodological. Critics argue that contingent valuation techniques do not yield reliable results. Defenders argue that the techniques are adequate or improving. They argue that if surveys were conducted more carefully, or with certain controls, then inconsistencies and intransitivities would disappear.[39]

---

[38] See Eric A. Posner, The Strategic Basis of Principled Behavior: A Critique of the Incommensurability Thesis, 146 U. Pa. L. Rev. 1185 (1998).

[39] See the essays in Valuing Environmental Preferences: Theory and Practice of Contingent Valuation in the US, EU, and Developing Countries (Ian J. Bateman & Kenneth G. Willis eds. 1999); Using Surveys to Value Public Goods: The Contingent Valuation Method (Robert Cameron Mitchell & Richard T. Carson eds. 1989).

However, as several economists have acknowledged, the main problem with contingent valuation of environmental goods is conceptual, not methodological.[40]

When people are asked for existence values, they often respond in strange ways. They refuse to answer surveys on environmental goods or, by way of protest, register a zero valuation or unrealistically high valuations that agencies must ignore. They provide valuations that are invariant across large and small parcels of wilderness or quantities of wildlife or that are inconsistent or intransitive. Their answers depend on the order in which questions are asked and are sensitive to the wording of the questions.[41]

Although one might hope for improvements that will eliminate inconsistencies, it is clear that no amount of methodological refinement will eliminate protest responses in the form of unrealistically high or low valuations or refusals to answer. But these responses must reflect something. When people give valuations of zero or infinity, these responses should be interpreted as an assertion that the question does not make sense, not as an assertion that the respondent would give nothing to save the environmental good (in the first case), or everything (in the second case), or is acting strategically (in either case) since he must know that his response will be disregarded. This is even more clearly true when people refuse to respond to the surveys.

What about reasonable CVs for environmental goods when people do not directly enjoy them? When people have no direct experience of environmental goods and claim to be willing to pay just for their existence, then (aside from the option value for possible use) this CV cannot reflect the goods' contribution to well-being. The dollar amounts in the survey responses should be interpreted as a valuation of the violation of a moral commitment, not as a valuation of an environmental amenity. To the extent the dollar amounts thus reflect moral commitments, aggregating them does not give one information about the effect of the project upon overall well-being or upon other moral criteria.

Why should agencies not pay attention to valuations people attach to violations of moral commitments? To see why, suppose that P believes that he will gain $100 from the construction of a dam as a result of lower electricity bills, but believes that construction of a dam is immoral. Q would lose $125

---

[40] See, for example, Charles R. Plott, Contingent Valuation: A View of the Conference and Associated Research, in Contingent Valuation: A Critical Assessment 470–73 (Jerry A. Hausman ed. 1993); Donald H. Rosenthal & Robert H. Nelson, Why Existence Values Should *Not* Be Used in Cost-Benefit Analysis, 11 J. Pol'y Analysis & Mgmt. 116 (1992). But compare Gardner M. Brown, Jr., Economics of Natural Resource Damage Assessment: A Critique, in Valuing Natural Assets (Raymond J. Kopp & V. Kerry Smith eds. 1993).

[41] See Plott, *supra* note 40, at 471–73, for a brief summary of the literature.

as a result of higher fish prices, but believes that construction of a dam is a moral obligation ("the march of progress"). P would pay an additional $50 to see his moral commitment vindicated, and Q would pay an additional $100 to see her moral commitment vindicated. These additional payments are, we suggest, morally irrelevant; they neither change the effect of the project on overall well-being (which is negative) nor change the moral status of the project in some other regard. The project is either moral or immoral or morally controversial; its morality does not depend on how much people are willing to pay to vindicate their moral views. If P receives a large inheritance, and so is willing to pay another $200 to see the project stopped, and this increase reflects his moral values, then it is false to say that the moral status of the project switches back from positive to negative, just because P's inheritance now enables him to outbid Q. Now it might be the case that an agency should take account of people's disinterested moral views when it decides whether to implement a project. We discuss this possibility below. The point to understand is that even if an agency should, it should not try to monetize these disinterested preferences, and should not include them as part of a CBA.

How then should environmental commitments (as well as other disinterested preferences) be recognized? There are two possibilities. The first is that agencies should have the minimal task of determining the effect of projects on overall well-being, and different political actors—Congress, the courts—should enforce moral commitments. No one asks agencies to decide whether to implement such "projects" as abortion legalization and capital punishment. In the political arena controversial issues are resolved not by cost-benefit analysis, but instead by political and moral debate through which people find common ground. Congress and the courts can erect constraints that bind agencies. These constraints prevent agencies from approving projects that, among other things, involve racial discrimination, the use of fetal tissue, and experimentation on people who do not give their consent—regardless of the extent to which the benefits of these projects outweigh their costs.[42]

The second possibility is to allow agencies to take into account the full range of moral considerations bearing upon projects and not simply the criterion of overall well-being. We have no objection to this alternative, in theory, as long as it is understood that the agency should not resolve a question of fairness, or deontological rights, or distributive justice—say, the use of fetal tissue in medical research—by engaging in cost-benefit analysis. The use of fetal tissue might be morally correct, or wrong, or morally con-

---

[42] See Department of Transportation, *supra* note 28.

troversial, but the resolution does not depend on whether one side of the debate is willing to pay more than the other side of the debate, in order to see its views embodied in the law. In any event, we suspect that agencies should not be given this authority, because they are not generally well positioned to make moral decisions other than decisions regarding welfare maximization.

Finally, it bears emphasis that the case of morally motivated preferences is arguably just the most extreme example of disinterested (non-welfare-relevant) preferences. Consider the case where P prefers the project not because it matters to his own well-being, but also not because he takes it to be morally obligatory; rather, P's preference is motivated by some kind of group loyalty. Or consider the case where P prefers the project out of a sense of obligation to his children. If we had a full and persuasive theory of how preferences should be restricted—of how the distinction between disinterested and welfare-relevant preferences should be drawn—then we might suggest that agencies should ignore *any* kind of disinterested preference or preference-motivating consideration in calculating WEs. But no such theory is yet at hand. The case of morally motivated preferences is, for now, the only case in which it is (1) clear that CVs can be disinterested and therefore can diverge from WEs and (2) practicable to enjoin agencies that they should adjust CVs so as to eliminate the effect of disinterested preferences.[43]

B.   Lack of Information

Suppose that an agency is considering whether to construct a park. A person P has no knowledge about the advantages of parks, so his CV may diverge from his WE. Specifically, let us imagine that P's CV is −$10: when asked how much he would pay or require to be paid in the project (park) world so as to be indifferent between the park and the status quo, he answers that he would require to be paid $10.

Further, let us define P's CV-I as his hypothetical willingness to pay, given complete information: the amount that, if P were fully informed, he would pay or require to be paid in the project (park)[44] world so as to be

---

[43] Compare David Sobel, Well-Being as the Object of Moral Consideration, 14 Econ. & Phil. 249 (1998), which provides a critical overview of philosophical attempts to distinguish between disinterested and welfare-relevant preferences.

[44] For the remainder of this article, unless otherwise noted, it should be assumed that all payments to or from persons are made in the project world. This is a basic feature of the CV. That construct—by contrast with the so-called equivalent variation, or EV—is defined in terms of the amount that persons would pay or require to be paid in the project world, not the status quo world. Similarly, all the variations on the concept of CV that we shall develop here, such as the WE, or the CV-I, shall be based on project world payments. Each of these variations can be matched with an analogous variation on the concept of the EV; but since

indifferent between the park and the status quo. Assume that CV-I is $100. (Note that CV-I, not CV, is the number that an agency elicits when respondents to valuation surveys are provided with detailed information about projects, such as in the Grand Canyon study.) Which number should the agency use in its CBA of the park? In particular, what is P's WE? Is it −$10, $100, or something else?

The answer depends on the nature of the project. Let us distinguish several possibilities.

First, people might costlessly and rapidly acquire the relevant information when the project is completed. P's CV is low (−$10) because he thinks that park views are ugly; in fact, they fill him with joy. His belief is wrong because he has never seen a well-maintained park. If P held the correct belief, he would be willing to pay $100, his CV-I; and if the park is implemented, his actual valuation will change to $100. Under these conditions WE is equal to CV-I rather than CV, and P's valuation should be treated as though it were $100.[45]

Second, people might never acquire the relevant information. P might undervalue the park, because he thinks it is full of common and easily cultivated plants, when in fact they are rare and difficult to grow—something P never learns. So P's actual valuation of the park remains −$10, even after the park is implemented. This is one type of case where correcting for preference distortion by looking to undistorted preferences (CV-I) turns out to be a mistake. Although CV-I is $100, P's WE cannot be greater than $0. As we have elsewhere argued in greater detail, it is a necessary condition for a person to benefit from a project that (at some point) she prefer that project to the status quo; actual preferences have at least *that* role on the correct theory of well-being, however idealized. This premise, a quite basic one, has the immediate implication that P cannot benefit from the project just described—in other words, that his WE must be $0 or negative.

A harder question is whether P's WE should be equated here with his CV (−$10), or set at 0, or given yet another (negative) value. That question

---

cost-benefit analysts in practice use CVs, not EVs, our focus will be upon the CV and its refinements.

[45] As we explained in Section I, we do not have a general theory as to how fully informed preferences and uninformed preferences interact, to produce well-being. But we are reasonably convinced that, in the particular case at hand, WE is equal to CV-I rather than CV. What if a period of time must pass during which P learns to appreciate parks? The simple solution is to discount to present value the amount he would pay once informed. P's CV should not be treated as though it were $100 but as though it were $100 discounted to present value. However, this simple solution is merely a suggestion on our part; the propriety of discounting future benefits and harms is a controversial issue that we do not have space to discuss here. See generally Richard L. Revesz, Environmental Regulation, Cost-Benefit Analysis, and the Discounting of Human Lives, 99 Colum. L. Rev. 941 (1999).

cannot be answered without a particular view as to how uninformed and fully informed preferences interact to produce well-being. On one view, WE is properly $-\$10$ because it is P's uninformed preferences that constitute (and continue to constitute) P's view of the world. On another view, both types of preferences have a coequal role in shaping welfare; P cannot be made better off by a project that he never comes to actually prefer, but neither can he be harmed by a project that, if fully informed, he would prefer. Thus he neither truly gains, nor truly loses, and his WE is neither his CV nor his CV-I but $\$0$. We confess to being swayed by the view that equates WE and CV (in the case at hand), but it is beyond the scope of this paper to take a definitive stand on that.

Several variations on this second case should be mentioned. One is where CV and CV-I have the same sign, but differ in amount. P barely likes the park, and would like it more if he knew about it; or he detests it, and would slightly dislike it if better informed. Note that the premise we invoked above—actual preference is a necessary condition for benefit—does not here help in deciding where WE lies in the range between CV and CV-I. A different variation on the second case is where information changes P's behavior. For example, a well-informed P would go into the park (and CV-I is based on the prediction that he would thus behave), but in fact P never learns something significant about the park and never goes in. In this particular example, P is not made better off by the park because he does not use it; his WE is not greater than $\$0$, even though CV-I is $\$100$. Generalizing from the example is hazardous, but it at least seems clear that WE can differ from CV-I by virtue of the behavioral impact of imperfect information.

Finally, the second case can be varied by having the information change instrumental judgments rather than intrinsic valuations. The project is not a park, but an air quality project, which P values (and would value under full information) in light of the effect of air quality on his longevity. He believes, incorrectly, that the cleanser used by the project would actually reduce his longevity, so CV $= -\$10$. In fact, the cleanser improves air quality, so CV-I $= \$100$. Here, the case for setting WE equal to CV-I even if P remains uninformed about the project once implemented seems stronger.

To turn now to a third case: people might acquire the relevant information only if the agency feeds it to them. The agency might have to distribute leaflets describing the benefits of parks or invest in television commercials. These activities are costly and should be included in the cost of the project. If the cost of disseminating information is high enough that the project has negative value once that cost is taken into account, then we have the second case, above. Otherwise, the agency performs an educative function as well as implementing the project.

C. Objectively Bad Preferences

Suppose some people support or oppose projects because their actual as well as fully informed preferences are sadistic. A person might favor a park because he wants to see a neighbor's beloved home demolished in order to make way for the park, or he may oppose the park because he does not want his neighbors to benefit from higher property values. Or a person might oppose AIDS research because he dislikes homosexuals and drug users with whom he is acquainted and does not want to see their suffering relieved. Or, suppose some person prefers a way of life that is clearly worthless. Person P wants to spend his days in an opium-induced haze, rather than working, developing relationships, accomplishing intellectual or practical goals, starting a family, or doing anything else; and this preference is wholehearted, in the sense that it conflicts with no second-order preference of P's[46] and does not change under full information.

The sadism and drug fixation[47] examples show that objective criteria are sometimes plausibly relevant to agency decisions and CBA.[48] The simplistic account of their relevance runs as follows: If the status quo is bad for person P in light of objective criteria, and the project is better for P in light of objective criteria, then the move from the status quo to the project improves P's welfare, regardless of P's actual preferences. But this simplistic account is incorrect, given the role of actual preference as a constituent (if not the sole constituent) of well-being. If the project is objectively better for P than the status quo, but P prefers (and would continue to prefer) the status quo, then his WE for the project cannot be larger than 0. This is parallel to the point we made in Section IIIA, about the size of WE given a deviation between actual and fully informed preferences.

So how should WE be calculated where criteria of objective welfare value and actual preferences diverge? In theory, we could think of the interaction between preference and value, to produce WE, along the following lines: Define a new measure, CV-O, as the amount that P would be willing

---

[46] See Harry G. Frankfurt, Freedom of the Will and the Concept of a Person, 68 J. Phil. 5 (1971) (discussing second-order preferences).

[47] The opium user's case is a case of drug fixation, not addiction, since his preference for drug use is fully informed, wholehearted, and (let us assume) otherwise autonomous. Government might plausibly prohibit the fully autonomous use of drugs just because that activity is thought to be objectively worthless.

[48] For arguments to the effect that objective values are a component of welfare, see, for example, John Finnis, Natural Law and Natural Rights (1980); Martha C. Nussbaum, Nature, Function, and Capability: Aristotle on Political Distribution 145 (Oxford Studies in Ancient Philosophy, supp. vol., Julia Annas & Robert Grimm eds. 1988); George Sher, Beyond Neutrality: Perfectionism and Politics (1997).

to pay or be paid for the project if his preferences perfectly tracked considerations of objective value; and then set WE depending upon the divergence between CV-O and CV. For example, P is willing to pay $10 for a new art museum, but if his preferences tracked objective values, he would be willing to pay $50; WE would be some amount between $10 and $50, depending on how exactly actual preferences and objective values interact under the right theory of well-being.

But this suggestion is not practicable. Calculating CV-O would be a hopelessly difficult task for agencies that were sincerely trying to measure WEs. Further, instructing agencies to determine CV-O would significantly increase the opacity of CBA and thus the extent to which agencies can pursue their own agendas rather than sincerely attempting to perform CBA. We are less sanguine about integrating considerations of objective value into agency decision making than we are about integrating the informational considerations discussed in Section IIIA—what P would prefer under full information, as against what he actually prefers—given that claims of objective value are highly contestable and not amenable to empirical testing.

A simpler and more practicable approach would be for objective values to bear upon agency valuations like this: If P is so perverse as to prefer a project that is clearly[49] objectively bad, or to disprefer a project that is clearly objectively good, then P's WE should be taken by the agency to be $0. Otherwise (except in the case of changing preferences, to be considered momentarily), objective values should be ignored.[50]

A plausible supplement to this proposal is to use objective values as a tool for choosing between different sets of actual preferences, where those preferences change over time. Suppose that P prefers the status quo but would prefer the project if it were implemented. Specifically, suppose that P's CV for the project based upon status quo world preferences is −$20, while his CV for the project based upon his project world preferences is $15.[51] In this sort of case, textbook CBA has no resources for assigning P

[49] "Clear" is an attempt to sort between value choices sufficiently uncontestable not to give rise to serious problems of transparency, and so on, and other sorts of value choices. It might not work.

[50] It is plausible that objective values are one component of well-being, but—as we stated in Section I—we are unsure whether this plausible suggestion is really correct. Thus the recommendation advanced in the text is a contingent one: If the designer of administrative procedures takes objective values to be a component of well-being, then he should instruct agencies to calculate WEs as recommended. The recommended use of objective values as a tiebreaker for choosing between different sets of actual preferences, set forth immediately below, is similarly contingent.

[51] This distinction between status quo world and project world preferences should not be confused with the quite separate distinction between CVs and EVs. (As we have already explained, textbook CBA generally uses CVs, not EVs, and the analysis in this paper is similarly focused upon CVs.) The CV/EV distinction concerns where dollar payments occur: P's

a unique project valuation. Both the status quo world and the project world preferences are actual-preference sets of P; given the view of welfare traditionally associated with cost-benefit analysis, namely, the unalloyed actual-preference view, there are no grounds for picking one set or the other as the unique basis by which to measure P's valuation of the project. Thus the textbook approach (at least in theory) would be to calculate two aggregate CV measures for the project described here, one assigning P a CV of −$20 and the other assigning her a CV of $15. This could, in turn, produce an indeterminacy at the level of aggregate CV and, thus, in the overall cost-benefit evaluation of the project. For example, if the project described here affects no one else in the world but P and Q, and Q's CV for the project is −$10, then one aggregate CV measure of the project (−$20 + −$10 = −$30) counts it as worse than the status quo, while another aggregate CV measure of the project ($15 + −$10 = $5) counts it as better.

However, a more sophisticated account of welfare may enable us to choose between conflicting sets of actual preferences and, thus, to assign a unique WE to persons whose preferences change over time. If P's status quo preferences are distorted in some way, while his project world preferences are undistorted, then plausibly P should be assigned a unique WE based upon his undistorted, project world preferences. This is, in effect, what we argued above where the distortion at hand was P's lack of information: where the project would increase P's information, such that his (uninformed) status quo preferences and (informed) project world preferences are different, agencies should use the latter in calculating P's WE. The same is plausibly true where the distortion at hand is the deviation between P's preferences and criteria of objective value. If the project is clearly better for P than the status quo, and if P's project world preferences favor the project while his status quo preferences do not, then the agency does have a basis for choosing between the two sets of preferences, and it is willingness to pay based upon project world preferences that the agency should use to calculate a unique WE for P.

D. Adaptation

Suppose people are not willing to pay for parks because they have adapted to a world without parks. Or they have persuaded themselves that

---

CV is the hypothetical dollar payment in the project world that would counterbalance the project's effect on him, while his EV is the hypothetical dollar payment in the status quo world that would counterbalance the project's effect on him. If P has unchanging preferences, then his EV and CV can still differ, but his CV will be a single, unique amount. By contrast, if P has different preferences in the project and status quo world, then this can give rise to two different values for his CV. That is our concern here.

only rich people need parks, because rich people are effete and weak. Merely informing people about the benefits of parks, then, will not cause people to change their preferences. Indeed, to keep the example as clear as possible, we will assume that people are well informed.

One should distinguish two kinds of adaptation. In the first case, people's adaptive preferences never change. Whether or not the agency creates a park, P will always oppose parks because of his impoverished childhood. For the same reason that agencies should not implement projects that benefit people only if they obtain information that will forever elude them, agencies should not implement projects only because the projects are ranked higher by idealized ("nonadaptive") preferences[52] that will never become actual. Where P prefers and continues to prefer the status quo (given his actual, adaptive preferences), the project cannot be welfare improving for him—his WE for the project cannot be greater than $0—even if P's idealized, nonadaptive preferences point in favor of the project.

To see why, imagine that the overburdened housewife discussed by Sen[53] opposes a project to create a well near her home, preferring the long walk to the river (which, let us assume, the well project will make inaccessible to her). If her preferences truly cannot be expected to change in response to the project—if she will use the well with regret, continuing to prefer the world in which she walked to the river—then it is hard to say how the project would make her better off. The project might be supplemented with educational efforts, in the hope that the housewife will develop different preferences as a result of education. But in the limiting case where the housewife's preferences are irrevocably entrenched, by virtue of upbringing, her WE for the project is no greater than $0.

As with the information case, the issue remains whether the housewife should be counted as being hurt by the project or instead be given a WE of $0. If the answer is $0, then the housewife is in effect ignored by the agency in its cost-benefit analysis of the well project even though her actual valuation of that project is negative. If the answer is to set WE equal to CV, then the agency would simply not take account of the fact that the housewife's preferences are adaptive. This might be the right answer, because the third alternative—to choose a number between $0 and the CV that properly reflects the degree of adaptiveness—is unpalatable. The problem with this alternative is its excessive difficulty. The problem here is even more difficult than the problems posed by lack of information. One can

---

[52] Everything we say in this section holds true regardless of the specifics of a theory for deciding when preference are inappropriately adaptive and for constructing idealized, nonadaptive preferences.

[53] See note 22 *supra*.

more easily imaginatively construct informed preferences out of uninformed preferences than nonadaptive preferences out of adaptive preferences. If the housewife prefers the walk because she was abused as a child, can we imagine what her preferences would be if she had not been abused? If she prefers the walk because she has unconsciously absorbed the views of her neighbors, can we imagine what her preferences would be if she has not unconsciously absorbed these views? We doubt that these questions can be answered, and we are sure that administrative agencies are in a poor position to answer them.

One possible solution to the problem is to use criteria of objective value. This is the solution we tentatively recommend. Sometimes adaptive preferences will also be clearly objectively bad. Perhaps this is true of the well case: perhaps it is clearly objectively better for the housewife to use the well than to walk to the river. If so, our suggested rule for the case of objective value would come into play,[54] and the housewife's WE would be set at $0. Sometimes, however, a person P can have an adaptive preference for the status quo—that preference can be rooted in some injustice or other deficiency of the person's background—even though the status quo is not objectively bad, relative to the project.[55] Imagine that P prefers one kind of recreation to another, that P continues to do so under full information, and that this recreation is not objectively harmful for P, but that P would prefer the second recreation had not the first been the only recreation available to him during an impoverished childhood. In this sort of case, we suggest, WE should be set equal to CV.

In the second variation on the case of adaptive preference, people's preferences change over time. P has an actual preference for the status quo, rooted in some unfortunate feature of his background, while his nonadaptive or idealized preference is in favor of the project. But P's actual preference is not fixed; if the project were implemented, he would come to prefer it. So P's CV based upon status quo world preferences is, say, −$55, while his CV based upon project world preferences is, say, $25. In this sort of case, at least in theory, criteria of adaptiveness could be used for choosing between P's conflicting sets of actual preferences, even where the project and the status quo are objectively fine and thus criteria of objective value provide no basis for making the choice. The agency could assign P a unique WE equal to $25 on the grounds that his actual project world preferences are nonadaptive while his actual status quo preferences are adaptive. But

---

[54] See Section IV C.

[55] Compare L. W. Sumner, Welfare, Happiness and Ethics 156–71 (1996) (distinguishing between the view that preferences must be objectively good to be welfare constitutive and the view that preferences must be autonomous to be welfare constitutive).

given the especial malleability of the concept of adaptive preferences (even as compared to the notion of an objectively "good" or "bad" way of life, which seems to have more commonsense resonance), we are quite skeptical that agencies should be instructed to use adaptiveness as a tiebreaker when status quo and project world preferences are different.

To sum up: Adaptiveness is a separate way in which preferences can be distorted, distinct from the lack of information and the lack of objective value. P's preference for the status quo can be adaptive, even if this preference is fully informed and even if the status quo is not clearly objectively worse than the project. Thus, in theory, adaptiveness could be a separate basis for agencies to reject CVs and a separate ingredient in the calculation of WEs. However, this suggestion strikes us as impracticable. In some cases, P's adaptive preference will also be uninformed, or objectively bad, or both, and in such a case a WE will appropriately be calculated for P based on considerations of full information or objective value. But adaptiveness per se should not, we think, be a component of agency decision making. If P's preference is distorted solely because it is adaptive, and in no other way, then agencies should stick to textbook CBA and use P's CV as the measure of the project's welfare impact upon him.

E. Wealth Distortions

Cost-benefit analysis is inaccurate, quite apart from the divergence between CVs and WEs, by virtue of the fact that both measures reduce welfare impacts on project winners and losers to dollars—which are, in turn, differentially productive of welfare in different persons. In particular, the marginal utility (strictly, marginal increase in overall well-being) of a dollar expended by a person poor in total wealth is generally larger than the marginal utility of a dollar expended by a rich person. The apparent solution is for agencies to weight WEs by the marginal utility of money. There are technical problems in constructing the right weighting factor; there may also be problems in transparency and reliability. But quite apart from technical and implementation difficulties, the proposal that agencies correct for the distorting effect of endowments by weighting WEs raises the subtler issues of (1) market adjustment (where rich and poor effectively undo the project picked out by weighted CBA, since their behavior is driven by unweighted preferences);[56] (2) the possible welfare superiority of money transfers to agency projects; and (3) the redistributive objection, which points out that weighted WEs (or CVs) would cause agencies to seek out

---

[56] Similar phenomena are widespread and much discussed—for example, the possibility that safety regulations cause consumers to take less care. See, for example, W. Kip Viscusi, Fatal Tradeoffs: Public and Private Responsibilities for Risk 223–27 (1992).

redistributive projects rather than projects that solve market failures. These are the issues we will focus upon here.

Suppose that an agency must decide whether to construct a park in a wealthy neighborhood or a poor neighborhood.[57] All people have the same preferences, which include a desire for more park space. More people live in the poor neighborhood than in the wealthy neighborhood, but the wealthy people are willing to pay more for the park because of their lower marginal utility of money. The agency, however, adjusts CVs using the marginal utility of money and relies upon the adjusted CVs to justify placement of the park in the poor neighborhood.

The result of the agency's action is that the property values in the poor neighborhood will rise relative to the property values in the rich neighborhood. If the poor people are renters, landlords might terminate their leases, convert to condominiums, and sell the condominiums to rich people. The rich people move out of the old rich neighborhood, and the poor people move out of the old poor neighborhood. Perhaps they simply switch places. Far from benefiting poor people, the agency's project benefits landlords, who are likely to be relatively wealthy, while causing a large welfare loss as people engage in unnecessary migration. It would be better to give the park to the rich neighborhood.[58]

This is an extreme version of what might happen. Another possibility is that some poor people own their own homes. These people would benefit from the increase in property values. But they would presumably sell to rich people, so the result of the agency decision to place the park in the poor rather than rich neighborhood is a transfer of wealth from rich people and poor people who do not own their homes to poor people who do own their homes. The rich people lose because they must pay to move to a new neighborhood, likewise, the poor people without homes. The poor people with homes gain because of the increase in property values.

In this second case, the project of building the park in the poor neighborhood is welfare inferior to the project of building the park in the rich neighborhood plus arranging a lump-sum payment from the rich people to the

---

[57] So the two outcomes being compared by CBA are "park in wealthy neighborhood" and "park in poor neighborhood." One would then be designated the status quo world, the other the project world, for purposes of calculating CVs. The discussion that follows does not depend upon which outcome, "park in wealthy neighborhood" or "park in poor neighborhood," is in fact designated as status quo.

How is the park funded? The funding scheme (for example, a scheme for funding the park through taxes) might well affect the welfare of persons in the "park in wealthy neighborhood" world, as compared to the "park in poor neighborhood" world. But we ignore this complication.

[58] See Sunstein, *supra* note 23, at 283.

poor people (or poor people with homes), which would not necessitate the uprooting and migration of large populations. However, agencies do not have the authority to order lump-sum payments. The agency's choice, then, is between benefiting some poor people through a project that diminishes the welfare of other poor people and a few rich people or benefiting a few rich people and some poor people at the expense of other poor people. Both are welfare inferior to the option of building the park in the rich neighborhood plus a transfer.

Suppose finally that people cannot move and that, therefore, property values do not adjust. The rich people will stay in the rich neighborhood and the poor people will stay in the poor neighborhood. Suppose that the rich people's aggregate CV is $1,000 and the poor people's aggregate CV is $500. If the park is placed in the poor people's neighborhood, they will obtain a value of $500. And this is true even of the very poor people who do not own their homes. One might argue that a better project would be the construction of the park in the rich neighborhood with a transfer, of say $600, from rich to poor.[59]

What should the agency do in these three cases? In the first case, by virtue of market adjustment, placing the park in the rich neighborhood turns out to be welfare superior to placing the park in the poor neighborhood. Here, it seems straightforward to us that the park should be placed in the rich neighborhood.[60] In evaluating the project, the agency should take account of market adjustment; if market adjustment undermines the value of a superficially attractive project, the project should not be implemented. Because placing the park in the poor neighborhood does not, after market adjustment, increase overall well-being, the welfare criterion counts against the park's placement there.

The second and third cases are more difficult. The fact that a third project—placing the park in the rich neighborhood and transferring wealth from rich to the poor—is welfare superior to the projects under consideration hardly seems relevant when that third project is not available to the agency. We know of no agency in the U.S. government that has the authority to order wealth transfers, and there are many good reasons for denying them this authority.

---

[59] See Louis Kaplow & Steven Shavell, Why the Legal System Is Less Efficient than the Income Tax in Redistributing Income, 23 J. Legal Stud. 667 (1994). Note, however, that an even better project than that might be constructing the park in the poor neighborhood and transferring $600 or a larger amount from rich to poor.

[60] This seems straightforward to us, insofar as the agency is solely concerned with overall well-being. If the agency is instructed to consider both overall well-being and other moral criteria, it might have grounds for placing the park in the poor neighborhood, although, in the particular example at hand, it is hard to see what those grounds could be.

Assume that this does not change and that in the second and third cases the only options available to the agency are the simple ones. What should the agency do? There is a plausible case that the agency should just choose the option identified by marginal utility weighted CBA, namely, placing the park in the poor neighborhood.[61] After all, this is welfare superior to placing the park in the rich neighborhood; weighted CBA is accurate, to this extent. To be sure, it might be the case that welfare-improving transfers through the tax and welfare system are not made because Congress has other things on its mind, and not because the optimal distribution of wealth has been achieved. But the agency has identified a way of increasing overall well-being and should implement it, and if this result is welfare inferior to an alternative that is politically impossible, that is irrelevant.

One might object to the approach suggested here because it would give agencies a license to search for projects that have differential impacts on rich and poor, and then approve them not because of their public good aspects, or because they satisfy other standard market failure rationales for government intervention,[62] but because of their redistributive effects. Agencies will implement projects in poor neighborhoods until people in those neighborhoods are as rich as people in rich neighborhoods. Weighted CBA gives agencies a license to undertake projects that increase overall well-being just by changing the distribution of wealth. This is the redistributive objection to weighted CBA.

The redistributive objection is a cogent one. There may be good, principled reasons against authorizing agencies to approve projects that are welfare increasing merely in virtue of their redistributive effects—because such a practice produces a disincentive to the accumulation of wealth, thereby decreasing overall well-being in the long run; or because, at some point, it invades the property rights of the rich; or perhaps for some other reason. But if this conclusion is correct, then the appropriate response of agencies is not to return to conventional CBA, with unweighted CVs (or WEs). There is no reason to think that a simple cost-benefit comparison of a project and the status quo, which uses unweighted CVs (or WEs), should generally reach the same result as a sophisticated cost-benefit comparison of a project and the status quo, which uses weighted CVs (or WEs) and also integrates market adjustment effects and incentive effects into the outcomes being compared.

---

[61] For purposes of the discussion, we shall assume that in the second case the option of placing the park in the poor neighborhood is indeed superior, using weighted CBA (and taking account of market adjustment), to the option of placing the park in the rich neighborhood.

[62] See, for example, Stephen G. Breyer, Regulation and Its Reform (1982) (presenting market failure rationales); Anthony Ogus, Regulation: Legal Form and Economic Theory (1994) (same).

Rather, we see two possible responses to the redistributive objection. First, as we suggested above, in theory agencies should count as costs perverse incentives that might result where CVs are weighted to correct for the declining marginal utility of wealth, that is, in inverse proportion to total wealth, thus producing a long-run disincentive to the accumulation of wealth. They should also include costs that would result from market adjustments, like the migration between neighborhoods discussed in the example above. Without consideration of market adjustments and perverse incentives, the placement of the park in the poor neighborhood looks better (using weighted CBA) than the placement in the rich neighborhood. With consideration of market adjustments but not perverse incentives, the placement of the park in the poor neighborhood looks better (using weighted CBA) than the placement in the rich neighborhood if rich and poor cannot move (the third case described above) but not necessarily if they can (the first and second cases). Finally, with consideration of both market adjustments and perverse incentives, the placement of the park in the poor neighborhood may not look better (using weighted CBA) than the placement of the park in the rich neighborhood even in the case where rich and poor cannot move.

To see how perverse incentives might be factored into agency decision making, consider once more the case of immobile neighbors and assume the following. Rational Rick, in response to the park's placement in the rich neighborhood, will pursue one employment plan, E1. In response to the park's placement in the poor neighborhood, he will pursue another plan, E2, that earns him less wealth (because he has less incentive to earn wealth if he thinks that the park agencies and other agencies will weight his CVs in inverse proportion to his wealth). Assume further that Rick's CV for E1, as against E2, is $200. Then $200, as weighted for Rick's marginal utility of wealth, could be added to the rich persons' CVs ($1,000, again appropriately weighted) in determining the monetized benefits of placing the park in the rich neighborhood.

A second possible response to the redistributive objection is for agencies to rely upon a rule of thumb. The problem with the first suggestion is that it may be impossible to calculate the ancillary costs—work disincentives, market adjustment—created by a project and by the particular technique that the agency uses to evaluate it. This is similar to the problem of calculating objective values. A possible rule of thumb is that agencies should avoid projects that produce large and concentrated losses, even on wealthier segments of society, and that provide minimal benefits, even to poorer segments of society. The EPA might rely on such a rule of thumb,[63] although

---

[63] See True, *supra* note 33, at 324.

it has not given its reasons, and one might suspect that the EPA avoids such projects simply because their high visibility makes them politically dangerous. Nevertheless, there is a rationale for this rule of thumb, and that rationale is that as losses associated with the project decline, and as the number of people affected declines, the magnitude of the project's perverse effect on work incentives and the danger of market adjustment also decline.

IV. SOME INSTITUTIONAL CONSIDERATIONS

Administrative agencies make decisions within a political structure and have important political purposes. They are agents of Congress, the president, and the people. If they are not supervised, they may regulate in a way that does not serve the public interest. The capture theory of regulation, according to which regulated industries bend agencies to their will, identifies one kind of agency problem. More broadly, an agency might serve the interests of its administrator, or of its bureaucracy, or of influential citizens or groups, rather than the interests of hierarchical superiors in the political branches or the general public. This danger might lead Congress, the president, or the courts to supervise agencies very closely. But if agencies are supervised too closely, then the various advantages that flow from specialization and division of labor are lost. Some balance must be struck between deference and supervision.

One technique for supervising agents (and agencies) is to require them to disclose information about their behavior. This is the political advantage of cost-benefit analysis: it forces agencies to be clear about the basis of their decisions, and this facilitates monitoring by other actors. Agencies that engage in direct welfare evaluation of projects may get the evaluation right or wrong, but it is very difficult for other actors to evaluate the agencies unless the agencies quantify or monetize costs and benefits. Analysis of some agency actions in fact suggests that agencies do not get it right. Agencies that do not use cost-benefit analysis make inconsistent assumptions about valuation of life.[64] Even agencies that use cost-benefit analysis make inconsistent assumptions about valuation of life, discount rate, and no doubt much else (see the Appendix), but it is easy to identify these inconsistencies and ask the agencies to justify them.[65] If an agency assumes a high valuation of life when justifying a regulation that injures one industry, while assuming a low valuation of life when rejecting a regulation that injures an-

[64] See Viscusi, *supra* note 57, at 264; John F. Morrall III, A Review of the Record, 10 Regulation 13, 30 (1986).

[65] See U.S. General Accounting Office, *supra* note 29.

other industry, and the regulations are in other respects identical, suspicions will be aroused that the second industry has captured the agency.[66]

Appendix Tables A1 and A2 show that agencies have tremendous freedom in choosing discount rates and valuations of life, despite efforts by OMB to impose some order.[67] An agency can apparently use a discount rate of 0.03 and a valuation of life of $5.8 million in one regulation and a discount rate of 0.1 and a valuation of life of $1.5 million in another regulation. That means that the agency could assert that a regulation that, say, saves 10 lives per year for 5 years produces benefits of as much as $266 million, or as little as $57 million. Cost-benefit analysis constrains agencies, but not as much as its proponents might hope.

Given that CBA currently imposes little constraint on agencies—at least, when the regulation will produce statistical deaths over a long period of time—one might hesitate about giving agencies more freedom by allowing them to correct for preference distortions. If agencies are permitted to modify CVs in order to account for distortions in preferences, then (1) inconsistencies in assumptions will be difficult to identify, and indeed (2) agencies may be able to conceal improper goals. This is a serious possibility, but the harm can be limited if agencies are required to explain deviations from CV baselines and use uniform adjustments. If an agency adjusts CVs to account for wealth differences, for example, it should be required to explain what the wealth differences are and what weightings are used to make the adjustments. The agency's argument for the adjustment will have to be reasonable, and, more important, it will have to be consistent across regulations, so that the agency cannot opportunistically change assumptions in order to justify some regulations and not others.

V. Conclusion: Some Proposals for Reform

Cost-benefit analysis poses a number of difficult conceptual problems. We have discussed five; these are the treatment of disinterested preferences, of preferences distorted by lack of information, adaptation, and objective badness, and of preferences whose influence is inflated or deflated by relative wealth. If agencies did not face information and decision costs and were not subject to political constraints, they could maximize overall well-being in several straightforward ways:

[66] This is an interpretation of Corrosion Proof Fittings v. EPA, 947 F.2d 1201, 1222–23 (5th Cir. 1991), which criticized the EPA for defending a regulation on the basis of a valuation for lives saved that is higher than that used to reject other regulations.

[67] See Office of Management and Budget, Benefit-Cost Analysis of Federal Programs: Guidelines and Discounts, 57 Fed. Reg. 53,519 (1992).

1. They would ignore disinterested preferences or else treat widespread moral commitments as (nonmonetized) constraints on projects.
2. They would use informed preferences when persons will become informed as a result of the project or when preferences are instrumental rather than intrinsic; they would otherwise rely, at least to some extent, on uninformed preferences; and they would consider information dissemination as a potential supplement to the project, with its own benefits and costs.
3. They would discount adaptive preferences, and rely to some extent on idealized, nonadaptive preferences in calculating WEs—if, for example, such preferences will become actual as a result of the project and even, perhaps, if adaptive preferences are entrenched.
4. They would ignore objectively bad preferences.
5. They would adjust for wealth distortions by weighting for marginal utility.

But agencies do face information and decision costs. Such fallible agencies must use whatever decision procedure minimizes the sum of these costs and the cost of error. Although it is hard to generalize, a number of comments can be made.

Even fallible agencies can successfully ignore objectively bad preferences when preferences violate widespread, uncontroversial intuitions about valuable and worthless behavior. Moreover, fallible agencies can successfully ignore disinterested preferences in certain situations: they should not use existence values for environmental entities. However, the appropriate response is not always so straightforward. A person might have a high CV for, say, a public commuter train both because of a self-interested preference for convenient transportation and because of a disinterested preference for environmentally sound transportation. A person might have a high CV for a bridge both because it reduces his cost of transportation and because it annoys his neighbors. These CVs would be reflected in market behavior as well as in survey results. In such cases of mixed preferences, the ideal agency would sort them out, but a real agency probably cannot. The real agency might plausibly choose to rely on traditional CVs on the theory that self-interested preferences tend to have much greater influence on CVs, except in domains where it seems likely that disinterested preferences dominate. The most important such domains are ones where the actual CVs are low, because in such cases the direct impact of the project is small and moral feelings are relatively powerful. Environmental regulation is such a domain, and that is why existence values should be ignored.

Real agencies are, we suspect, unlikely to be able to distinguish adaptive preferences from nonadaptive preferences. Accordingly, we think that agen-

cies should ignore this category. It is likely that most extreme cases are better handled as objectively bad preferences. For example, preferences of drug addicts, whether or not adaptive, are generally considered objectively bad. Indeed, the fact that agencies would ignore the preferences of nonaddicted drug users suggests that objective value is the more appropriate category.[68]

When preferences are uninformed, agencies should sometimes make adjustments. If projects that are based on informed preferences actually improve well-being, either because people do not need information in order to receive the benefit or because they are likely to obtain information after the project is implemented, then there is a good case for constructing informed preferences. However, when uninformed people do not become informed as a result of the project, matters are more complex, as we discussed in Section III.

We are more optimistic about restricting preferences on the basis of information than on the basis of adaptation for two reasons. First, it is easier for agencies to derive informed preferences from uninformed preferences, than nonadaptive preferences from adaptive preferences. One can, for example, compare the behavior of people who are informed about nutrition and people who are uninformed about nutrition; one can observe changes in behavior as people obtain information; and so on. But it is, even as a conceptual matter, hard to distinguish adaptive preferences from nonadaptive preferences. Many influences contribute to the formation of preferences; distinguishing "corrupt" from "pure" influences may be impossible or even meaningless. Second, it is easier to give people information than it is to change their preferences. Indeed, people will seek out information because it can help them satisfy their desires, but people commonly resist efforts to change their preferences. The first requires education; the second requires brainwashing.

Finally, when CVs are distorted because of wealth differences, real agencies might use distributive weights and then make further corrections to deal with the problems of market adjustment and perverse incentives. We have outlined how these corrections could be made. But the more practicable course, we think, is for agencies to rely upon the rule of thumb suggested above—to avoid projects that have a large impacts upon the distribution of wealth.

Further, it is unclear whether the basic idea of distributive weighting is itself a feasible one. It may be just too complicated for Congress or OMB to specify a methodology for weighting CVs, in inverse proportion to total

[68] See note 47 *supra*.

wealth or income, that is reasonably accurate (in compensating for the declining marginal utility of wealth) and that agencies can use with reasonable success. The feasibility of distributive weighting has been much debated, without a clear resolution, by welfare economists. If distributive weighting is not feasible, then agencies should probably use unadjusted CVs when the distribution of wealth among the winners does not differ much from the distribution among the losers. When the distributions differ greatly, an agency could refrain from implementing the project. One possible alternative route would be to inform Congress and hierarchical superiors in the executive branch; these officials might be willing to arrange for compensation of the losers or some other politically desirable outcome. Finally, an agency could use a procedure other than CBA for comparing the project and status quo[69]—in effect, a procedure that reaches the same kind of results as weighted CBA but is more feasible—but this would lead back to the problem of market adjustment and perverse incentives. We are skeptical that agencies can really take account of these.

A comment about this last point should be added. Supporters of CBA have traditionally argued that it avoids distributional judgments and allows agencies to focus on efficiency improvements that their expertise puts them in a position to identify. Critics of CBA have pointed out that distributional judgments cannot be avoided. The efficiency of a project is a function of its distributive effects. Unadjusted CVs are unacceptable because they reward people on the basis of wealth, yet wealthy people are on average likely to value a dollar on the margin less than poor people are. Properly adjusted CVs would result in possibly massive redistribution to the poor as agencies implemented projects that tax the rich (because their marginal dollars are worth little to them) and benefit the poor (who can then use valuable marginal dollars for other purposes). Our view is that in theory agencies should take account of the costs of market adjustments and work disincentives, and if they could do this, then properly adjusted CBA would not result in a massive redistribution of wealth. In practice, agencies are unlikely to be able to calculate these costs, so certain broad constraints—against projects that have large impacts on the distribution of wealth, for example—might be justified. This may be a rough description of agency practice, but there is much room for improvement.

---

[69] See Adler & Posner, *supra* note 1, at 229–33 (discussing multidimensional assessment).

APPENDIX

TABLE A1

VALUATIONS OF LIFE

Regulation	Value ($Million)
Department of Transportation, Federal Aviation Administration, Proposed Establishment of the Harlingen Airport Radar Service Area, TX, 55 Fed. Reg. 32,064 (1990)	1.5
Department of Agriculture, Food Safety and Inspection Service, Pathogen Reduction: Hazard Analysis and Critical Control Point Systems, 61 Fed. Reg. 38,806 (1996)	1.6
Department of Health and Human Services, Food and Drug Administration, Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents, 61 Fed. Reg. 44,396 (1996)	2.5
Department of Transportation, Federal Aviation Administration, Aircraft Flight Simulator Use in Pilot Training, Testing, and Checking and at Training Centers, 61 Fed. Reg. 34,508 (1996)	2.7
Environmental Protection Agency, Protection of Stratospheric Ozone, 53 Fed. Reg. 30,566 (1988)	3.0
Department of Health and Human Services, Food and Drug Administration, Proposed Rules to Amend the Food Labeling Regulations, 56 Fed. Reg. 60,856 (1991)	3.0
Department of Transportation, Federal Aviation Administration, Financial Responsibility Requirements for Licensed Launch Activities, 61 Fed. Reg. 38,992 (1996)	3.0
Department of Agriculture, Food and Nutrition Service, Proposed National School Lunch Program and School Breakfast Program, 59 Fed. Reg. 30,218 (1994)	1.5, 3.0
Environmental Protection Agency, National Ambient Air Quality Standards for Particulate Matter, 62 Fed. Reg. 38,652 (1997)	4.8
Environmental Protection Agency, National Ambient Air Quality Standards for Ozone, 62 Fed. Reg. 38,856 (1996)	4.8
Department of Health and Human Services, Food and Drug Administration, Medical Devices: Current Good Manufacturing Practice, 61 Fed. Reg. 52,602 (1996)	5.0
Department of Health and Human Services, Public Health Service, Food and Drug Administration, Quality Mammography Standards, 62 Fed. Reg. 55,852 (1997)	5.0
Environmental Protection Agency, Requirements for Lead-Based Paint Activities in Target Housing and Child-Occupied Facilities, 61 Fed. Reg. 45,778 (1996)	5.5
Environmental Protection Agency, National Primary Drinking Water Regulations: Disinfectants and Disinfection Byproducts, 63 Fed. Reg. 69,390 (1998)	5.6
Environmental Protection Agency, Radon in Drinking Water Health Risk Reduction and Cost Analysis, 64 Fed. Reg. 9560 (1999)	5.8

TABLE A2

DISCOUNT RATES

Regulation	Discount Rate (Costs/Benefits)
Environmental Protection Agency:	
Emission Standards for Locomotives and Locomotive Engines (1997)	7/7
Requirements for Lead-Based Paint (1996)	3/3
Nuclear Regulatory Commission, License Term for Medical Use Licenses (1997)	7/not quantified
Occupational Safety and Health Administration:	
Respirator Protection (1998)	7/not quantified
Indoor Air Quality (1994)	10/not quantified
Food and Drug Administration:	
Food Labeling Regulations (1993)	5/not quantified
Regulations Restricting the Sale and Distribution of Cigarettes and Smokeless Tobacco to Protect Children and Adolescents (1996)	Not quantified/3
Federal Aviation Administration, Aircraft Operator Security (1997)	7/7
National Highway Transportation Safety Administration, Federal Motor Vehicle Safety Standards; Lamps, Reflective Devices and Associated Equipment (1997)	2–10/2–10
Consumer Product Safety Commission, Requirements for Labeling of Retail Containers of Charcoal (1996)	5, 10/5, 10

SOURCE.—Edward R. Morrison, Judicial Review of Discount Rates Used in Regulatory Cost-Benefit Analysis, 65 U. Chi. L. Rev. 1333, 1364–69 (1998).