PAPERS on Economics & Evolution

MAX-PLANCK-GESELLSCHAFT

1124

Why Satisfy Preferences?

by

Daniel M. Hausman

The *Papers on Economics and Evolution* are edited by the Evolutionary Economics Group, MPI Jena. For editorial correspondence, please contact: evopapers@econ.mpg.de

ISSN 1430-4716

Max Planck Institute of Economics
Evolutionary Economics Group
Kahlaische Str. 10
07745 Jena, Germany
Fax: ++49-3641-686868

Why Satisfy Preferences?

Daniel M. Hausman

University of Wisconsin-Madison*

Abstract

Contemporary mainstream normative economists assess policies in terms of their capacities to satisfy preferences, though most would concede that other factors such as freedom, rights, and justice are also relevant. Why should policy be responsive to preferences? This essay argues that the best reason is that people's preferences are in some circumstances good evidence of what will benefit them. When those circumstances do not obtain and preferences are not good evidence of welfare, there is little reason to satisfy preferences.

* * *

It is hard to assess social policies, in part because there are many relevant moral considerations. Policies should be efficient and fair (both in their execution and in their consequences). They should respect people's rights. They should be responsive to the objectives and commitments of the population. They should respect obligations to future generations and perhaps to the earth itself. There are multiple moral criteria in terms of which to assess policies. These are difficult to balance and in some cases controversial, and they are hard to define precisely.

In this essay I shall argue that a technique developed by welfare economists – cost-benefit analysis – can help assess policies, but only under limited conditions. In my view, "net benefit" – the economist's measure of potential improvements in "efficiency" – sometimes indicates potential improvements in social welfare and in that case it is relevant to policy assessment. As most welfare economists would emphasize, net benefit is rarely the only important factor. Other considerations are typically relevant and often just as important.

This essay has five sections. Section 1 discusses the general problems of evaluating the efficiency of policies and lays out the basic elements of cost-benefit analysis. In section 2 defends what I take to be the best answer to the question my title poses. It argues that preference satisfaction is relevant to policy assessment as an indicator of welfare. Section 3 defends a cautious conception of cost-benefit analysis that is justified by the link between preference satisfaction and welfare delineated in section 2. Section 4 considers whether preference satisfaction bears in other ways on policy and criticizes applications of cost-benefit analysis that cannot be justified by the connections between preference satisfaction and welfare defended in section 2. Section 5 concludes.

1 Policy evaluation, efficiency and cost-benefit analysis

"Efficiency" is only one consideration in policy assessment, and it may not be the most important, though it is usually relevant and apparently uncontroversial. Agreement on the importance of efficiency may however disguise disagreement concerning what efficiency is. In the most general sense, efficiency is not wasting means to an end. Disagreements about the end to be sought and about what are the means that should not be wasted may thus imply different conclusions concerning efficiency. In this essay, I shall

* This develops further views expressed in Chapter 8 of *Preference, Value, Choice and Welfare* (2011).

assume, as economists do, that the goal is individual welfare or well-being (which I will treat as synonymous) and the means are scarce resources.[1] More efficient policies achieve or make it possible to achieve more welfare with the same resources (Legrand 1991). Efficiency is only one important consideration in policy assessment. The most efficient policy may not be the best policy.

It is not easy to determine which policy is most efficient at promoting welfare, let alone which is best overall. The efficiency of policies at promoting welfare depends on their consequences, which are uncertain and hard to predict. Even when there is scientific consensus concerning the consequences of policies, as there is concerning the effects of green-house gas emissions, powerful private interests can confuse the public and bias policy-makers. Political deliberation is not impartial and it is not fully rational. (Indeed at least in the U.S. one sometimes wonders whether it might not be entirely irrational.) Furthermore, even when the physical consequences of policies are known and accepted, it is difficult to translate those into consequences concerning welfare. Policies bear on human welfare in many ways. (Is a new missile system likely to improve overall welfare more than an expansion of drug-treatment centers or high-speed rail lines?) The costs and benefits of policies are also spread out over time. Should policy-makers discount future costs and benefits, and if so, at what rates? Policies may even influence which people exist. How should the welfare of possible future people count? Even if policy-makers understood how states of affairs bear on the interests of everyone affected, it is not clear how policy makers should compare the welfare of different people, as they must if they are to determine net gains or losses. It is unclear even what welfare is. In addition to all these problems, policy-makers must ask whether other and possibly better alternatives have been overlooked. Where did the list of "serious" alternatives come from?

These are just the main complications that remain when policy-makers temporarily set aside all ethical considerations that ought to guide their deliberations, apart from efficiently pursuing the well-being of the population. Rationally setting a course through this thicket is not easy, and policy-makers are bound to seek either a short cut or some way around the problems. If only there were a method of assessing the efficiency of policies that cut through these difficulties!

On one interpretation, which I shall criticize, cost benefit analysis may appear to be just such a method. In my view, cost-benefit analysis rests on six assumptions:

(1) Individual welfare can be measured by preference satisfaction.

(2) Social welfare depends (a) on *the social capacity to satisfy preferences* – that is, on the size of the economic pie – and (b) on *the distribution of that capacity* – that is, on how the pie is sliced.

(3) Mainstream normative economists[2] are concerned exclusively with efficiency, which they identify with (a), the social capacity to satisfy preferences. Questions of distribution are for others to address. Accordingly, economists take there to be an unambiguous increase in welfare in adopting some policy x if *and only if* somebody prefers x to the status quo and nobody prefers the status quo to x. In that case, x is "Pareto superior" to the status quo or a "Pareto improvement." Economists thus avoid making interpersonal comparisons of welfare or claims about the net

[1] It is not obvious that policy should always aim to increase welfare, since policies that improve the welfare of the malicious, envious, or malevolent may be objectionable regardless of their consequences for the welfare of others. But this essay takes for granted the popular view that concerns about welfare should be one of the most important factors governing policy. It then focuses on the question of when preferences can be taken as a guide to welfare.

[2] There are many differences among normative economists, and the claims I am making about them are not true of figures such as Amartya Sen (1977, 1987, 2002) or those libertarian economists who are concerned with freedom rather than welfare. I shall be discussing only orthodox mainstream welfare economists. In terms of numbers and influence, they dominate the field.

welfare gains or losses when (as is of course typical) a policy benefits some people and harms others.[3]

(4) Economists measure preferences by how much people are willing to pay for alternatives, and they infer willingness-to-pay from market behavior or, more controversially, from surveys.

(5) Economists take there to be a "net-benefit" in adopting some policy x, when the "winners" are willing to pay more to bring about *x* than the amount that the "losers" would require in compensation in order to be indifferent to the adoption of policy *x*.[4]

(6) Economists take a net benefit to imply an increase in efficiency, because it is in principle possible (regardless of whether it is feasible) to transfer funds from the winners to the losers to transform a policy that provides a net benefit into one that is a Pareto improvement over the status quo. A policy that by means of transfers can in principle be transformed into a Pareto improvement is a "potential" Pareto improvement. In other words, economists take net benefit to indicate an enlargement of the economic "pie", though whether there is an actual improvement in welfare also depends on how the pie is sliced.

In fact there are serious problems with several of these assumptions. Contrary to assumption 4, people's willingness to pay depends on how much income and wealth people have, not just on their preferences. So the net benefit implied by measures of willingness to pay is not a measure exclusively of capacity to satisfy preferences. By merely redistributing income or wealth one might be able to make the status quo into a strict Pareto improvement over some proposed policy x, even though x is a potential Pareto improvement over the status quo. As Tibor Scitovsky (1941) and Paul Samuelson (1950) showed, it may be the case that x is a potential Pareto improvement over the status quo, and the status quo is a potential Pareto improvement over x. That possibility conclusively refutes assumption 6. The fact that some policy is a potential Pareto improvement does not imply that it enlarges the economic pie. But for the purposes of this essay, I shall set aside these problems with assumptions 4 and 6. My concern is instead with the connection between preferences and welfare.

Cost-benefit analysis computes the net benefits or net losses of policies in terms of people's willingness to pay. Net benefits are taken to indicate an increased capacity to satisfy preferences and thus to quantify the efficiency (in this special sense) of alternative policies. On the interpretation that I shall criticize, the so-called *ex ante* view (which relatively few economists would whole-heartedly espouse), cost-benefit analysis permits the policy maker to evade almost all the problems of assessing the efficiency of policies that were listed at the beginning of this section. All policy makers need to do is to commission economists to measure the net benefits of policies – that is, the sum of what members of the society would be willing to pay to implement a policy minus the amount needed to compensate those who do not like the policy. Assessing the efficiency of social policies is not then exactly easy, because estimating net benefits from people's market behavior or with the help of surveys is a complicated technical business. But the task can be left to economists. Policy-makers do not need to employ scientists to determine the physical consequences of policies, and they can avoid the political pressures that lead to biased

[3] Consider, for example, a situation where everybody needs more food, but some people are badly malnourished, while other people have almost all that they need. A distribution that allowed widespread death and malnutrition would surely lead to less overall welfare than one that avoided either. But if one ignores the possibility that food distribution affects productivity and hence future consumption possibilities, then none of the distributions that avoid either satiation or wasting food is a Pareto improvement over any other. They are all equally efficient in the narrow sense of efficiency employed by welfare economists. The claim that there is an unambiguous increase in welfare only if there is a Pareto improvement permits economists to be agnostic about whether actual redistributions ever increase welfare.

[4] I am describing what is known as "compensating variations." One could instead measure net benefit by means of so-called "equivalent variations" – that is, asking those who do not want the policy how much they would pay to prevent its being implemented and asking those who do want the policy how much they would have to be paid to be indifferent to implementing the policy.

assessments. They need take no stand on discounting, nor concerning what welfare is, nor how the consequences of policies bear on welfare. They still need to be concerned, lest a superior alternative be overlooked, but determining which is most efficient among the policies under consideration apparently requires nothing more than measuring net benefit, which can be left to economists.

How is it possible for economists to spare the policy maker from having to cope with all the difficulties: determining the consequences of policies and how those translate into quantities of welfare, figuring out what welfare is and how to make interpersonal welfare comparisons, deciding how to discount gains and losses of welfare that occur at different times? How can all these difficult technical and ethical problems be transformed into the hum-drum difficulties of measuring willingness to pay? Obviously, they cannot be. So we need to look more deeply into cost-benefit analysis to see what is wrong with this extreme interpretation and whether any other interpretation is tenable. A great deal is at stake, because cost-benefit analysis is, in practice, the only available quantitative guide to policy making, and contentious and potentially misleading numbers may be better than no numbers at all. (Then again they may not be.) We need to look more closely.

2 Welfare and preference satisfaction

Without a theory of welfare, how is it possible to make claims concerning the efficiency of policies at providing welfare? If it is not possible, then economists must be committed to some theory of welfare, and on one reading of welfare economics, they appear to be proponents of the view that welfare is constituted by preference satisfaction: The preference-satisfaction theory of welfare hold that the fact that Jack prefers *x* to *y* makes it the case that *x* is better for Jack than *y*.

On this interpretation, welfare economics rests on a gross error. It is obviously not the case that Jack's preferring *x* to *y* makes it the case that *x* is better for him than *y*. Jack might have false beliefs about the characteristics and consequences of *x* and *y*. *x* and *y* might have nothing to do with Jack. Jack might not have heard of either *x* or *y* and have no preferences among them, yet *x* might nevertheless be better or worse for him than *y*.

These are just the most obvious difficulties. Setting them aside, and setting aside other difficulties concerning past preferences, preference changes, intransitive preferences, conflicts between preferences and meta-preferences, and preference distortions owing to cognitive flaws, there is something odd about the notion that Jack's preferences could determine whether *x* or *y* is better for him. Suppose that *x* and *y* are just the same states of affairs except that in *x* Jack gets an A in his calculus class and a B in his history class, while in *y*, Jack gets a B in calculus and an A in history. According to a preference-satisfaction view of welfare, *x* is better for Jack than *y* if and only if Jack prefers the state of affairs where he gets an A in calculus and a B in history to one where he gets a B in calculus and an A in history. How could his preference determine whether *x* is better for him than *y* (Kraut 2007)?

One quick answer is that Jack will be happier if the state of affairs that he prefers obtains. But this answer supposes that welfare is happiness, not preference satisfaction. Furthermore, knowing that he got an A in calculus is what makes Jack happy, not just getting the A. If he is shipwrecked on a deserted island before learning his grades and never comes to know them, then *x* would not make him happier than *y*. In virtue of satisfying his preference, would getting an A in calculus still make Jack better off?

A second quick answer fares no better. Perhaps, one should look to Jack's *reasons* for preferring *x* to *y*. Perhaps his grade in calculus is much more important to Jack, because he wants to go to graduate school in physics. On the assumption that Jack has a reasonable idea of what is likely to benefit him, *x*, the alternative he prefers is accordingly better for him. But this answer supposes that there is some other

notion of welfare or benefit than preference satisfaction upon which *x* and *y* bear. *x* is not better for Jack than *y* because he prefers it. On the contrary, he prefers it because he perceives that it is better for him.

A third answer may seem more promising. It could be the case that a career as a physics teacher, to which Jack aspires, is better for him than being a history teacher for no reason other than the fact that he wants to be a physics teacher rather than a history teacher. If Jack's career aspirations were different, his performance in his history class might be more important than his performance in math, and it would accordingly be the case both that he would prefer *y* to *x* and that *y* would be better for him than *x*. Though it would not strictly-speaking be the case that Jack's preference *among x and y* would determine which is better, other preferences of his – namely his career aspirations – would determine which is better for him.

This response conflates (a) successfully pursuing a worthwhile objective (such as a career aspiration) with (b) satisfying preferences. Though Jack's plan to become a physics teacher depends on his preferences, plans are more than preferences. In addition to satisfying or frustrating his preferences, his performance in calculus influences his prospects of succeeding in a presumably worthwhile project; and it is in virtue of its contribution to his project that *x* is better for Jack than *y*. So this third attempt to explain how it could be that preferring something makes it better for you winds up, like the first two attempts, denying what it aimed to explain. Succeeding in worthwhile projects rather than satisfying preferences makes states of affairs good for people.

It should thus be no surprise that there are so many apparent counterexamples to the view that welfare is preference satisfaction. Nor should one be tempted to follow philosophers who have argued that welfare is constituted by the satisfaction of rational, well-informed, and self-directed preferences (Gauthier 1986; Goodin 1986). It is mysterious how the fact that Jill's preference is rational, well-informed, and self-directed helps to make it the case that her preference for *x* over *y* determines that *x* better for her. The only way that the fact that her preference for *x* over *y* is rational, well-informed, and self-directed matters is that such preferences are more likely to reflect a correct judgment on her part that *x* is in some other way better for her than is *y*.

But, as I have argued elsewhere (Hausman and McPherson 2009), perhaps economists are not committed to this indefensible view of welfare after all. Instead of maintaining that preferences determine welfare, they might assume that expectations concerning welfare shape preferences. More specifically, economists can proceed as follows. First, assume that people are good judges of the value of alternatives to themselves or at least better judges than others. For example, Jill knows that peanuts make her sick, while walnuts do not and for that reason she prefers walnuts to peanuts. Economists need not suppose that individuals are always good judges. Jill may not know that she also has an allergy to soft-shelled crabs and thus mistakenly express indifference between hard-shelled and soft-shelled crabs. But welfare economists can idealize, and in many contexts they can reasonably ignore mistaken preferences.

Second, economists typically assume that Jill's preference between *x* and *y* derives from her judgment that *x* is better for her than *y* – that is, that Jill is entirely self-interested. Obviously, people are not always thinking about themselves and unwilling to express any preferences among alternatives until they have some view about which will better promote their own welfare. But there is arguably a large domain of self-interested evaluation and choice.

Third, welfare economists can often ignore flaws in people's evaluative capacities and in the preferences that result. Psychologists and behavioral economists have documented a large number of these. If enrollment in retirement plans depends on whether participation is the default and one needs to opt out or whether not participating is the default and one needs to opt in, then it is hard to regard people's enrollment choices as reliable indicators of what is better for them. But economists may reasonably hold that such flaws and distortions frequently have only a small or unsystematic effect on people's preferences and choices.

When economists put these three assumptions (accurate beliefs, self-interest, and unflawed evaluation) together, they can conclude that Jill prefers *x* to *y* if and only if *x* better for her than *y*, not because her preference makes *x* better, but because her preference rests on her correct belief that *x* is better for her in some other way than her preferring it. The cost-benefit analyst can then take this third-person evidential connection between preference satisfaction and well-being as justifying welfare economics (Scanlon 1998, pp. 116-18).

This reasoning is illustrated in figure 1. Suppose there are only two alternatives, *x* and *y*, with *x* in fact better for Jill than *y*. *If* Jill judges correctly, then Jill ranks the expected benefit of *x* above that of *y*. *If* Jill is self-interested, then Jill prefers *x* to *y*. If Jill knows that she faces a choice between *x* or *y*, then she chooses *x*.

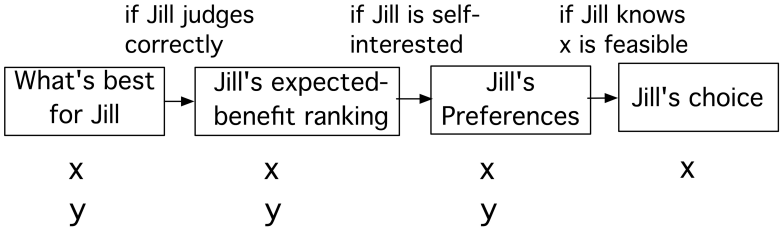

Figure 1

The economist can then work backward. From Jill's choice and the assumption that she her choice and preferences do not depend on cognitive flaws, the economist can infer Jill's preference – but only on the assumption that Jill knows that she could have chosen *y*. From Jill's preference for *x* over *y*, the economist can infer that Jill thinks that *x* is better for her than is *y* – but only on the assumption that Jill is self-interested. From Jill's judgment that *x* is better for her than is *y*, the economist can conclude that *x* is in fact better for her – but only on the assumption that Jill's judgment is correct. In this way welfare economists can draw inferences concerning well-being from people's choices without committing themselves to any theory of well-being. Though what people prefer does not *determine* what is good for them, it is sometimes *evidence* concerning what is good for them. As Elliott Sober pointed out to me, just as Socrates in Plato's *Euthyphro* said of piety that it is loved by the Gods because it is holy rather than holy because loved by the Gods, so I contend that an alternative is preferred by rational, self-interested, and well-informed people because it is good for them, not good for them because they prefer it.

Even in the contexts where people are not completely self-interested and in which their beliefs and deliberations are flawed, their preferences may be the best guide policy-makers have to what is good for them. Legislators typically know less of the circumstances of an individual such as Jill than she knows, and they probably have a less whole-hearted concern for her well-being than Jill does. The judgments of legislators about how to make Jill better off are often likely to be worse than her own judgments. It is also *safer* to rely on people's preferences. The mistakes individuals make about their own good may cancel out, and legislators will not arrogate to themselves the power to substitute their judgment for people's own judgments. Treating individuals as authoritative judges of their own interests is also one way of showing respect for them.

Nevertheless, *the fact that some alternative satisfies someone's preference does not make that alternative better for that person*. It is easy to forget this by sliding between thinking of utility as merely an indicator of preference and thinking of utility as pleasure or welfare. For example, if one describes Jack's charitable preferences by stating that Haitian earthquake relief is "an argument in Jack's utility function," then it seems to follow that giving relief to the victims of the earthquake "gives utility" to Jack – that is, makes Jack better off. But it only follows in virtue of an equivocation on the meaning of

"utility." Preferring *x* to *y* implies judging that *x* is better than *y*, *with respect to everything one takes to be relevant*, not judging that *x* is better *for oneself* than *y*.

People can distinguish between the questions, "Which is better, all things considered?" and "Which is better for me?" Having distinguished the two questions, it could be the case that the only considerations someone finds relevant are how *x* and *y* bear on his or her interests. Such complete self-interest is possible, though rare. What most people judge to be best and what they judge to be best for themselves will not always coincide, because most people think that other considerations than their own interests are relevant to their evaluations of alternatives. Most people have preferences among alternatives whose expected self-interested benefits are equal or unknown.

A cynical economist might reply, "Of course, people may tell themselves or others heart-warming stories about their selfless motivation, but in fact they always acts in what they believe to be their own interests. What explains why Jill regularly attends the Whooping Crane Lovers Association meetings are the pleasures she experiences from interacting with other members of the group, not whether whooping cranes live or die." Suppose this sketchy cynical explanation of Jill's behavior is correct. The fact that her preference for survival of whooping cranes motivates Jill to carry out actions that benefit her (whether or not she recognizes this) does not imply that the survival of whooping cranes itself benefits her. Nor the fact that acting on such a preference benefits her it imply that she believes that the survival of whooping cranes makes her better off. What purportedly explains her preference are the social and psychological benefits of acting on the preference not any benefits from the satisfaction of the preference. If the survival of the whooping crane were secured and Jill's preference thus satisfied, the Whooping Crane Lovers Association might collapse and her welfare diminish. Whether the self-interested benefits of hanging out with other environmentalists explains why Jill is devoted to protecting whooping cranes is irrelevant to how the fate of whooping cranes bears on Jill's well-being.

The evidential view of the connection between preference satisfaction and welfare does not support any version of a preference-satisfaction theory of well-being, and it has few implications for philosophical theories of welfare. It claims that, *regardless* of what philosophical theory of human well-being one accepts (other than an actual preference satisfaction view), preferences sometimes indicate well-being. If well-being turned out to be happiness, then the preferences of rational and self-interested individuals without mistaken beliefs would tell economists what makes them happy. If the well-being of individuals consisted of some set of objective goods, the preferences of those who are rational, self-interested and well-informed would tell economists what those sets contained.

One might question whether economists can avoid committing themselves to any theory of well-being. Unless economists suppose falsely that preferences are always self-interested, well-informed, and rational, they must rely on some idea of well-being to distinguish between self-interested and non-self-interested preferences and to decide when people are good judges of what in fact serves their interests. It might thus appear that economists cannot avoid the philosophical task of articulating a theory of well-being by adopting the evidential view of the relationship between preferences and welfare.

This objection points out correctly that economists need to know something about what is good for people in order to distinguish self-interested from non-self-interested preferences or to judge when people are good judges of what is good for them, but it does not follow that they need a philosophical *theory* of well-being for this purpose. For example, urban economists can easily see that a new community swimming pool contributes to individual welfare because they know that that people enjoy swimming and socializing at swimming pools. This conclusion assumes that enjoyment contributes to well-being. Similarly, economists can conclude that the prevalence of malaria-carrying mosquitos diminishes well being, if they assume that sickness and death diminish well-being. Platitudes concerning what makes people better or worse off like the claims that enjoyment contributes to well-being and illness diminishes it depend on no philosophical theory that specifies what things are intrinsically good for people and why.

Yet such platitudes are informative. Economists do not have to wait for a satisfactory philosophical theory of welfare before they can venture opinions concerning what makes people's lives better or worse. Welfare economists, like the rest of us, have beliefs about what makes people better off, even if philosophical modesty makes them reticent about trumpeting those beliefs. I would not be surprised if some of their beliefs are mistaken. For example, I suspect that many economists exaggerate the importance of material possessions in a good life. Be that suspicion as it may, economists know enough about the things that make people better or worse off that they can determine roughly when satisfying preferences enhances welfare.

In summary, when people's preferences are undistorted and largely self-interested and their beliefs are true, preference satisfaction indicates welfare. It is because preferences often indicate welfare that policy makers should aim to satisfy preferences.

3 Interpretations of cost-benefit analysis

The version of cost-benefit analysis sketched in section 1, which apparently transforms the complexities of assessing the efficiency of alternative policies into an exercise in preference measurement, uses willingness to pay *for alternative policies* as a guide to efficiency. This way of interpreting net benefit, the *ex ante* view, is, I believe, rarely defensible. The capacity of policies to satisfy preferences depends on the consequences of the policies and the extent to which those consequences satisfy preferences, not on *ex ante* preferences among policies. The *ex ante* view seems like an operationalization of the view that the social assessment of policies should depend on individual *assessments* of policies rather than a method for determining how to promote welfare. Since preferences among alternatives express *evaluations* of those alternatives,[5] which, as argued in section 2, need not be judgments about the expected self-interested benefits of alternatives, a social assessment of policies that relies on individual preferences among policies need not be an assessment of the extent to which policies promote welfare.

In its *ex ante* interpretation, net benefit can be a guide to welfare, but only in very special circumstances. Let P be the population whose willingness to pay for alternative policies is measured and P' be the population of those who are actually affected when the policies are implemented.[6] Willingness to pay *for policies* will be a guide to efficiency only if the willingness to pay of members of P is a reasonable indicator of what is good for the members of P'. When P and P' largely coincide and members of P know what the consequences of policies will be, net benefit on the *ex ante* interpretation may be a guide to efficiency. But these conditions are often not met. For example, suppose that a proposal to build a highway has, *ex ante* a large net benefit because people currently hate congestion and believe that the highway will relieve congestion. If the highway in fact increases congestion, it diminishes rather than increases welfare. If the current use of fossil fuels affects the welfare of people who will be born a century from now, and people today have no grasp of the circumstances or preferences of their distant descendants or simply do not care about them, then current preferences among policies will not be a good guide to which policy creates the greater capacity to satisfy the preferences of those affected by the policy.

On the *ex ante* interpretation of net benefit, in which the decision-maker can simply hand over to the economist the task of assessing the efficiency of policies by measuring willingness-to-pay, cost-benefit analysis in effect delegates to those whose willingness to pay is measured the difficulties the policy maker finds so challenging. In some circumstances – in particular, when the effects of alternative policies are broadly known and will not be long-lasting, when there is good reason to assume that preferences will remain stable, when preferences are self-interested and unlikely to be distorted – such delegation may be

[5] This claim is fleshed out and defended in my *Preference, Value, Choice and Welfare* (2011).
[6] Different policies might lead to different populations, but I shall ignore this complication.

reasonable. But *ex ante* net benefit is only rarely a good guide to which policies most enhance efficiency. When addressing policy choices such as whether to build nuclear reactors to generate electricity and where to locate them, it would be folly to rely on *ex ante* willingness to pay for the alternatives. Lay beliefs concerning consequences of nuclear power generation are confused and distorted, and the consequences of building such generators may persist for millenia. Considerations of respect for individuals may justify deferring to the values of the current population with respect to the consequences of policies *they* personally will experience, but what, if anything justifies deferring to preferences founded on mistaken or poorly supported beliefs or of allowing the values gleaned from surveys or market behavior to dictate how other people should be treated?

Fortunately, there is a more justifiable way to implement cost-benefit analysis. In particular, one can (as urged by those defending the so-called *ex post* interpretation of net benefit), examine willingness to pay for the expected *consequences* of alternative policies rather than willingness to pay for the policies themselves, where the expectations are grounded in the best available scientific predictions of the probabilities of the consequences. One still has the problem that the population whose willingness to pay economists measure need not be the same as the population that experiences the consequences of the policies, and if the policies change the way people evaluate alternatives, this problem will be serious. But at least the *ex post* interpretation avoids basing policy on popular misinformation or on the contingencies of people's present concerns for future people and consequences. Of course, this means that the decision-maker cannot dodge the question of deciding whose estimates of the consequences of policies to rely on and how to discount future benefits and costs. The dangers of corrupt science and biased political decision-making are back in force. But the *ex ante* approach only swept them under the rug.

The discussion in section 2 of the relationship between welfare and preferences makes clear that even if one accepts the *ex post* interpretation of net benefit, there are serious limits to the application of cost-benefit analysis. The fact that when people's preferences are self-interested and are not the result of cognitive flaws or false beliefs, then they are guides to their well being provides grounds for diffidence about overruling preferences and identifies sketchily the factors that determine whether it is reasonable to take preferences as indicating welfare. Practitioners of cost-benefit analysis who want to determine which policy has the greatest capacity to promote the welfare of the current population must take care that people's willingness to pay depends on true beliefs, self-interest, and undistorted cognition. These are demanding conditions, but there are circumstances in which, to a reasonable degree of approximation, they are met. For example, when deciding whether to build a bridge or a tunnel, it is reasonable to impute the welfare costs of the road noise from the bridge from the drop in the prices of houses located live near highways.

Sometimes satisfying individual preferences, which are themselves self-interested, well-informed, and free of any cognitive defects can be collectively self-defeating. At an exciting moment in a football game, each spectator might prefer to stand in order to see better. Yet when all stand, on average nobody sees any better and all are less comfortable. Problems like these infect actual cost-benefit analyses. Consider, for example, the Congressional Budget Office's cost-benefit analyses of increases in the gasoline tax and of increases in the average fuel economy demanded of new cars (Dinan and Austin 2004). These analyses quantify the welfare loss from driving smaller and less powerful cars by relying on market data that show how much people are willing to pay for larger and more powerful vehicles. The preferences implicit in car purchases are largely self-interested, and car buyers are reasonably well informed about the alternatives they face. No doubt there are distortions due to inertia, and people overestimate how much they will care about what car they drive. Advertisers also attempt to play upon longings for status and sex appeal that car ownership might magically produce. But it seems that preferences among cars are roughly the sort that cost-benefit analysis should rely on.

But since the advantages of a larger and faster car depend on the size and speed of other cars, reducing the size and acceleration of all cars could be like getting people at sports events to avoid

standing up to see better. If regulations or taxes make all cars smaller, someone who now needs a Hummer to be the king of the road could have the same experience in a SUV weighing half as much (Frank 2000). Car purchases plausibly reflect people's judgments of which cars will be best for them, *given the distribution of other cars on the road*. These preferences are not a good basis upon which to evaluate policies that will change that distribution. When the welfare effects of a policy on me depends in part on how that policy affects the behavior of others, then my willingness to pay, which economists infer from my market behavior will be a misleading guide to the consequences of the policy on my welfare, because my market behavior implicitly assumes that the behavior of others will not change. As Robert Frank has argued, a relational good that, given to one person, increases (or decreases) his or her welfare, may have little or no benefit (or cost) when given to many people.

Cost-benefit analysis is not a mechanical method of determining which policy is most capable of promoting welfare. Net benefit need not indicate what better satisfies preferences and preferences need not indicate welfare. Yet we have no other quantitative guide to policy-making, and policy makers need the information that cost-benefit analysis can sometimes provide. There are good reasons to continue to employ cost-benefit analysis, but economists must be aware of how tenuous an indicator of efficiency and welfare net benefit may be.

4 Non-self-interested preferences

Some economists have argued that policy makers should aim to satisfy preferences, regardless of whether those preferences are self-interested or based on justified beliefs. These economists must either believe that the satisfaction of preferences contributes to welfare, even if the preferences do not arise from justified expectations of self-interested benefits, or they must believe that satisfying preferences is an important social objective, regardless of whether it promotes welfare. For example, Vining and Weimer take cost-benefit analysis to provide "a framework for comprehensively taking account of the full range of social benefits and costs" (2010, p. 1; see also Zerbe *et al*. 2006). They argue that measures *of the net benefit* of policies should depend on the extent to which policies satisfy non-self-interested preferences such as distributional preferences.

> An alternative approach to measuring the social benefits of improvements to the circumstances of the least advantaged is based on the observation that many people derive utility from helping the disadvantaged. In other words, they are willing to pay something to help the most disadvantaged. The spontaneous outpouring of charitable giving after major disasters is one indication of such altruism. (Vining and Weimer 2010, p. 22).

To maintain that people "derive utility from helping the disadvantaged" could mean that people are made better off by helping the disadvantaged or that people prefer helping the disadvantaged to spending their money some other way. Willingness to pay to help the disadvantaged reveals the latter – that people prefer to spend some of their money on helping to other ways of using their funds. Willingness to pay to help the disadvantaged does not tell economists whether satisfying people's preferences for helping makes the helpers better off than would providing other goods or services. If helpers are altruists, as Vining and Weimer assert, then by definition they are not seeking their own advantage. An altruist prefers *x* over *y* because the altruist judges that *x* is better for other people than is *y*. *x* might also be better for the altruist. Self-interest and the interests of others are not necessarily opposed. But if *x* benefits the altruist, those benefits do not explain why the altruist prefers *x* to *y*. Though it is plausible that satisfying altruistic or malevolent preferences typically provides some benefit to those who hold such preferences, there is no general reason to believe that satisfying altruistic or malevolent preferences is better for the altruist or the malevolent person than lower ranked alternatives whose value derives entirely from self-interest.

Why then should satisfying non-self-interested preferences be counted among the "social benefits"

of policies. Welfarists who think that social policy should promote welfare have no reason to satisfy non-self-interested preferences. Welfarists want to know how policies affect people's welfare, not how much people who are not attempting to promote their own welfare would be willing to pay or require in compensation to institute one policy or another. To be sure, if altruistic preferences to help the disadvantaged are well-informed, then satisfying those preferences will most likely make the disadvantaged better off, but by the same token, satisfying well-informed malevolent preferences toward the disadvantaged will make them worse off. The welfarist is concerned with welfare. Preferences and willingness to pay are relevant only as indicators of welfare.

Vining and Weimer do not, however, appear to be welfarists. They are concerned with a broader notion of social value than welfare. But what social value is there in satisfying preferences when preferences are not evidence of well being? What justification would there be for a policy that satisfies preferences while at the same time harming people – as might be the case with respect to the preferences of the malevolent or when preferences are based on false beliefs? The most one could say is allowing individual preferences to influence policy, even when there is reason to believe that satisfying those preferences would not enhance welfare, is a way of showing respect for people or helping them to succeed in their projects. But should social policy respect all preferences and assist in all projects? Furthermore, when preferences are grounded in mistaken beliefs, satisfying them is not likely to help people to succeed in their projects.

Rather than taking net benefit to measure efficiency of policies – that is, the capacities of policies to promote welfare – and then to base policy on both efficiency and other ethical considerations such as equity, Vining and Weimer would make net benefit a measure of social value that takes into account efficiency in the promotion of welfare as well as all other moral views that citizens hold. But should policy assessment depend on how well policies actually measure up to moral demands, or should it depend on how well the population thinks policies measure up? In addition to examining the efficiency of alternative policies, should policy makers ask whether the consequences of the alternatives are equitable, whether they violate rights and whether they restrict freedoms, or should these moral considerations affect the assessment of policies via their influence on the overall preferences of the populace among policies and their consequences as reflected in data concerning willingness to pay? If one held that respect for people requires heeding their preferences, whether or not they are self-interested, well-informed or flawed or that the idea of popular sovereignty required that preferences rule, then one would have a reason to take preferences into account regardless of whether they indicated welfare or any social value. But these reasons to permit preferences to *rule* are not reasons to take preferences as determining social value. Though popular sovereignty may force a choice of x over y, it does not imply that x is the better choice.

In addition, it is questionable allowing net benefit calculated from surveys or market behavior to influence policy is a defensible way to operationalize the notion of popular sovereignty. People's house purchases have implications concerning how the location of an airport will affect their well being, but it is questionable whether such market choices are good evidence of people's all-things-considered evaluation of social policies. If asked to express a preference as a citizen among different locations for a new airport, individuals might want more information and a chance to deliberate. There is nothing exceptional or even unusual about an individual who buys an SUV and also supports legislation designed to make SUVs more expensive to purchase and drive. Popular sovereignty gives little reason to allow net benefit, when it is not an indicator of welfare, to influence policy.

As an indicator of welfare, net benefit can justifiably influence policy. When it is not an indicator of welfare, why measure net benefit or allow it to influence social policy? *Policy-makers have reason to satisfy preferences only if people have good reason to have them.* Representative governments in fact place a number of road blocks (in the form of legislative procedures, a bill of rights, and minority protections) in the way of responding immediately to non-self-interested preferences and ethical views of citizens.

In addition, willingness-to-pay for things such as distributive justice or environmental amenities, historical preservation, protection of endangered species or climate control is usually not a good indicator of how highly people value these things. Preferences are comparative evaluations, not the heavings of guts, and without having the feedback that comes with ordinary consumption nor extensive opportunities for thoughtful reflection on specific distributional questions, few people will be able to determine what their preferences are. When asked their willingness to pay, they will answer. Economists can also impute willingness to pay for redistributive policies from people's contributions to United Way. But the answers to surveys or the quantities imputed from charitable contributions will not reflect settled valuations of social policies.

The recognition that the point of welfare economics in general or cost-benefit analysis in particular is to use people's preferences as evidence to guide in the promotion of welfare and that the objective is *not* to satisfy preferences *per se* should lead economists to be suspicious about some applications of welfare economics. Consider, for example, attempts to measure the "non-use" or "passive-use" values of natural resources by means of contingent valuation, which in effect consists in asking people how much they are willing to pay to protect a marsh or an endangered species (Arrow *et al*. 1993). Even though it is reasonable in some contexts for economists to assume that people's preferences reflect their well-founded judgments of what will benefit them, following Sagoff (2004), I conjecture that willingness to pay to protect the environment is not one of those contexts. I do not believe that people's willingness to pay to protect the environment is mainly driven by their expectations of how much they personally will benefit. For example, even with the successful campaign to sow doubt concerning global warming, millions of older people are willing to pay a good deal in the form of higher energy costs to limit greenhouse gas emissions, even though they cannot reasonably believe that limiting greenhouse gases will benefit them personally. When economists measure the welfare effects of such policies by measuring how much people are willing to pay for them minus their costs, they get the wrong answer. Contingent valuation studies are generally erroneous measures of the impact of the environment on people's welfare (see also Sobel 1998, p. 251). Unless there is some reason other than the promotion of welfare why policies should be sensitive to people's preferences, contingent valuation studies of non-use value should not influence policy.

The conjecture I have offered made can be tested. Rather than asking people how much they would be willing to pay to protect whooping cranes or to preserve an old-growth forest, one could ask them to state the monetary equivalent of how much they expect personally to benefit. Instead of asking people, "How much would you be willing to pay to prevent whooping cranes from going extinct?" economists could ask them, "State the amount of money the loss of which would diminish your own welfare by the same amount that your welfare would be diminished if whooping cranes went extinct." If people's willingness to pay to prevent whooping cranes from going extinct expresses their judgment of how much the existence of whooping cranes contributes to their lives, the answers to the second questions should be as large as the answers to the first. I conjecture, in contrast, that the monetary estimates people would provide of the expected self-interested harm of the loss of a species or a forest will be much less than the amount they would be willing to pay to prevent the species from going extinct or the forest from being logged. In my view, which I think is shared by many, the extinction of species, especially of elegant and graceful creatures, is a bad thing that I am willing to help prevent, regardless of whether it makes me personally worse off. Well-executed surveys concerning estimates of monetary equivalents to welfare gains or losses more accurately elicit information concerning the welfare consequences of environmental preservation than contingent valuation studies, which overstate the welfare benefits to the current generation.

Implementing such surveys sensibly, like implementing a contingent valuation measurement, would be a complicated task. Figuring out how to estimate the welfare consequences to future generations and how much to weight them would be an additional problem. But once welfare economists recognize that

the relationship between willingness to pay, preferences, and welfare is evidential rather than constitutive, they should recognize that preference measurement, regardless of whether it take the form of measuring willingness to pay, is neither the only nor always the best way to gather evidence concerning the expected efficiency of alternative policies. When preferences are not self-interested and there are other ways to learn about people's expectations of benefit, economists should explore those other ways.

With respect to environmental protection and global warming, putting welfare estimation on a conceptually coherent basis could have political consequences that I would regret. An accurate measure of the welfare benefits of these policies would, I conjecture, show them to be considerably lower than current contingent valuation methods imply. Yet I personally favor strict environmental protections and severe limits on emissions of greenhouse gases. If welfarist policy makers were no longer led by contingent valuation studies to exaggerate the welfare gains of environment protection, they would favor fewer environmental protections and less stringent limits on greenhouse gas emissions.[7] The fact that the welfare benefits of environmental protections are, as I believe, less than what contingent valuation studies mistakenly indicate would not shake my support for such protections, because I believe that there are grounds apart from welfare to support them. But many people weight welfare more heavily than I do, and in this instance conceptual clarity could be politically harmful.

5 Conclusions

If, as I have argued, economists should take preference satisfaction only as evidence for welfare, then they should restrict the preferences from which they calculate net benefit to those that are self-interested. In the case of non-self-interested preferences they should attempt to measure expected benefits directly rather than attempting to infer them from measures of preferences. In addition, when preferences are self-interested and there is reason to measure them, economists should take care to rely on only undistorted preferences that rest upon accurate or at least well-justified beliefs.

Many practitioners of cost-benefit analysis in fact take steps to provide people with better information, to correct mistaken beliefs, and to avoid relying on distorted preferences, including preferences based on false beliefs (Adler and Posner 2006). Contemporary psychology has identified contexts in which people are likely to make mistakes, and policy analysts can use these findings to help decide whether to take people's preferences as guides to their welfare. What connects preferences and welfare are self-interest and people's judgments about what serves their interests.

When preferences are self-interested, well informed, and undistorted, they are a good guide to what benefits people. When these conditions are met, it is sensible for those seeking to promote welfare to employ methods of appraising policies such as cost-benefit analysis that rely on information concerning preference satisfaction. When these conditions are not met, economists can take steps to purify people's preferences of mistake and distortion so as to widen the domain in which these conditions are met, or they can measure expected benefit directly rather than inferring it from preferences.

In any case, efficiency and welfare are not the only relevant moral considerations. Though qualitative rather than quantitative, considerations such as justice, freedom, and rights are of great importance and can easily trump considerations of welfare. Cost-benefit analysis is currently the only source of quantitative information concerning the virtues and vices of alternative policies that policy makers have. But the numbers it provides are not always measures of the contributions that alternatives

[7] One might argue that with a low discount rate welfarists would have good reason to take strong actions to protect the environment. Whether this is the case depends heavily on how one decides to take into account the enormous uncertainties concerning the distant future.

are capable of making to welfare. Those who do cost-benefit analysis should employ it only when willingness to pay indicates preferences and preferences reflect welfare.

References

Adler, Matthew and Eric Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Cambridge, MA: Harvard University Press.

Arrow, Kenneth, Robert Solow, Paul Portney, Edward Leamer, Roy Radner, and Howard Schuman. 1993. "Report of the NOAA Panel on Contingent Valuation." *Federal Register* 58(10), 4601–4614.

Boardman, Anthony, David Greenberg, Aidan Vining, and David Weimer. 2010. *Cost-Benefit Analysis*. 4th ed. Englewood Cliffs: Prentice-Hall.

Broadway, Robin and Neil Bruce. 1984. *Welfare Economics*. Oxford: Basil Blackwell.

Dinan, Terry and David Austin. 2004. "Fuel Economy Standards versus a Gasoline Tax." *Congressional Budget Office*.

Fankhauser, Samuel, Richard Tol, and David Pearce. 1997. "The Aggregation of Climate Change Damages: A Welfare Theoretic Approach." *Environmental and Resource Economics* **10:** 249–266.

Frank, Robert. 2000. "Why Is Cost-Benefit Analysis So Controversial?" *Journal of Legal Studies* 29: 913-30.

Gauthier, David. 1986. *Morals by Agreement*. Oxford: Oxford University Press.

Goodin, Robert. 1985. "Laundering Preferences." In *Foundations of Social Choice Theory*, ed. J. Elster and A. Hylland, pp. 75–101. New York: Cambridge University Press.

Hammond, Peter. 1983. Ex-post optimality as a dynamically consistent objective for collective choice under uncertainty. In *Social choice and welfare*, ed. P. Pattanaik and M. Salles. Amsterdam: North-Holland.

Harris, R. and N. Olewiler. 1979. "The Welfare Economics of *ex post* Optimality." *Economica* 46: 137–47.

Hausman, Daniel. 2011. *Preference, Value, Choice and Welfare*. Cambridge: Cambridge University Press.

Hausman, Daniel and Michael McPherson. 1994. "Preference, Belief, and Welfare." *American Economic Review Papers and Proceedings* 84: 396-400.

_____. 2006. *Economic Analysis, Moral Philosophy, and Public Policy*. Cambridge: Cambridge University Press.

_____. 2009. "Preference Satisfaction and Welfare Economics," *Economics and Philosophy* 25 (2009): 1-25.

Hick, John. 1939. "The Foundations of Welfare Economics." *Economic Journal* 49: 696–712.

Kaldor, Nicholas. 1939. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *Economic Journal* 49: 549–52.

Kraut, Richard. 2007. *What Is Good and Why*. Cambridge, MA: Harvard University Press.

Layard, Richard and Stephen Glaister. 1994. "Introduction." in Layard, Richard and Stephen Glaister, eds. 1994a. *Cost-Benefit Analysis*. 2nd ed. Cambridge: Cambridge University Press, pp. 1-56.

Legrand, Julian. 1991. *Equity and Choice: An Essay in Economics and Applied Philosophy*. London: Harper-Collins.

Sagoff, Mark. 2004. *Price, Principle, and the Environment*. Cambridge: Cambridge University Press.

Samuelson, P. 1950. "Evaluation of National Income." *Oxford Economic Papers* (N.S.) 2: 1-29.

Scanlon, T.M. 1998. *What We Owe to Each Other*. Cambridge, MA: Harvard University Press.

Scitovsky, Tibor. 1941. "A Note on Welfare Propositions in Economics." *Review of Economic Studies* 9: 77-88.

Sen, Amartya. 1977. "Rational Fools: A Critique of the Behavioural Foundations of Economic Theory," *Philosophy and Public Affairs* 6: 317-44.

_____. 1987. *On Ethics and Economics*. Oxford: Blackwell.

_____. 2002. *Rationality and Freedom*. Cambridge, MA: Harvard University Press.

Sobel, 1998. "Well-Being as the Object of Moral Consideration." *Economics and Philosophy* 14: 249–81.

Sugden, Robert and Alan Williams. 1978. *The Principles of Practical Cost-benefit Analysis*. Oxford: Oxford University Press.

Vining, Aidan and Weimer, David L. (2010) "An Assessment of Important Issues Concerning the Application of Benefit-Cost Analysis to Social Policy," *Journal of Benefit-Cost Analysis* 1 (1): Article 6. Available at: http://www.bepress.com/jbca/vol1/iss1/6 DOI: 10.2202/2152-2812.1013.

Zerbe, Richard, Yoram Bauman, and Aaron Finkle. 2006. "An Aggregate Measure for Benefit–Cost Analysis." *Ecological Economics* 58: 449– 461.