

ARTICLE

Distribution and Disputation: Net Benefits, Equity, and Public Decision-Making

H. Spencer Banzhaf

North Carolina State University, Raleigh, NC, USA
NBER, Cambridge, MA, USA
PERC, Bozeman, MT, USA
Email: hsbanzh2@ncsu.edu

Keywords: benefit-cost analysis; distributional effects; history and methodology of economics

JEL classifications: B2; B4; D6; H5; Q5

Abstract

As its practitioners know well, benefit-cost analysis (BCA) walks a fine line between the positive and normative, between the science of economics and the art of political economy. Missteps threaten to undermine its credibility as a value-free science, while overcaution risks irrelevance to the pressing questions of the day. As BCA adapts to give more weight to distributional concerns, while operating in a more highly charged political environment than ever before, these tensions will only grow. For perspective, I reexamine three prominent episodes in the history of economics where these issues were vigorously debated: (i) The founding of the NBER by Wesley Clair Mitchell, who insisted that the organization eschew all policy recommendations; (ii) the introduction of the modern definition of economics as the study of tradeoffs by Lionel Robbins, who insisted welfare effects could never be aggregated; and (iii) the origins of BCA as a measure of income, which to first-generation practitioners seemed to foreclose the possibility of measuring “intangible” benefits like recreation opportunities, mortality risks, and equity. These episodes, together with critiques of economics from philosophers of science, suggest we are best served by being as transparent as possible about the ways values influence BCA reasoning, without arrogating political decisions into it.

[E]conomics is not ethics, though it borders on ethics; the line between them is a place where it behoves us to walk very delicately. It is now quite clear that neither the First nor the Second Welfare Economics have been walking delicately enough.
– Hicks (1975)

1. Introduction

Economists, more than anybody, should recognize the reality of tradeoffs. But, as human beings, we often want to have it both ways: On one hand, we want to be the policy advisor,

to speak to the pressing questions of the day, about how to increase the wealth of the nation, how to alleviate poverty and inequality. If only we could make people understand the long-term benefits of free trade, and so forth. On the other hand, economists aspire to rigorous science, uniting high theory with high-quality empirical work to test hypotheses and make causal inferences. And we understand Science to be objective and value-free. It focuses on facts, on what is true, rather than on what ought to be true.

These twin callings of scientific objectivity and policy relevance force practitioners of benefit-cost analysis (BCA) to walk a fine line between the positive and the normative, between the science of economics and the art of political economy.¹ Overcaution risks irrelevance. Clumsy recommendations undermine BCA's credibility.

Questions about distributional effects highlight these two dangers. One danger is to neglect distributional concerns as much as we traditionally have, making BCA appear irrelevant to the questions of the times. Perhaps for this reason, BCA seems poised to expand its scope to incorporate more analyses of distributional effects. For example, Germany incorporated them into its social cost of carbon (Astrid & Bünger, 2019), and the USA is currently considering revisions to BCA guidance that would encourage such analyses (US OMB, 2023). Reflecting these trends, the *Journal of Benefit-Cost Analysis* recently published a special issue on the topic.²

The opposite danger is to embed rash ethical judgments into our analyses about the ideal distribution of resources, making BCA appear unobjective. To cross the (admittedly fuzzy) line between economic science and ethics is to invite attacks on the entire enterprise. The attack on science, and its downstream effects on policy, is a matter of great concern these days to scientists and policymakers alike (e.g., McCarthy, 2019). While the scientific community is right to be concerned, those concerns should motivate us to reflect on the ways we may have invited some of these attacks, with clumsy rhetoric along the lines of "science says we should cut carbon emissions 50% within 10 years," or "science says we should wear masks and socially distance," and so forth.³ The real danger here is not that people will not believe us, but that they will! For notice the simple syllogistic logic for somebody who disagrees with the policy conclusion: (i) science says we should socially distance; (ii) I think it is wrong to socially distance; therefore, (iii) I must be against science. BCA faces similar risks.

Fortunately, these problems are not unprecedented. In this essay, I will review some of the ways economists have navigated them over the past 100 years or so, considering a few prominent historical episodes. These stories are comforting in themselves, for they remind us that we are not alone. But they also provide concrete lessons that will be useful going forward.

The roots of BCA's dilemma lie in the is-ought dichotomy. This logical dichotomy, especially in the realm of politics and policy, can be traced back at least as far Machiavelli.⁴

¹ Su and Colander (2013) discuss the distinction between the science of economics and the art of political economy, in the tradition of J. S. Mill and J. N. Keynes.

² See Acland and Greenberg (2023), Cecot (2023), Kniesner and Viscusi (2023), Lakdawalla and Phelps (2023), and Sunstein (2023).

³ Of course, many do recognize these risks. For example, surveying a heated dispute over a European Commission report on the risks of endocrine disruptors, Elliott and Resnik (2014) have issued a plea to scientists to be more transparent about their values.

⁴ In his advice to *The Prince*, Machiavelli wrote that "there is such a gap between how one lives and how one ought to live, that anyone who abandons what is done for what ought to be done learns his ruin rather than his preservation" (xv).

The dichotomy was developed most famously by David Hume, with his formulation that there is no ought from is (Hume's law).⁵ Hume based this supposed law, not on analytical relationships, but on the premise that our understanding of what is, of facts, comes only from sensory perceptions, and we can never literally *see* an ought. Closer to living memory, in the earlier 20th century, the logical positivists pushed the argument even further, arguing that only truth claims either provable by logic or verifiable empirically are cognitively meaningful, ruling out all ethics as not only unscientific but even as unmeaningful, except as mere expressions of preference.

As I discuss in Section 2, two important episodes from the history of economics in the first half of the 20th century addressed concerns about the fact-value dichotomy. One was the founding of the National Bureau of Economic Research (NBER) by Wesley Clair Mitchell for the scientific study of empirical economic relationships. The other is Lionel Robbins's critique of the interpersonal comparisons of utility as unscientific, the resulting fear of economics' policy irrelevance, and the search for a compromise in the potential Pareto test.

Section 3 introduces later critiques of the fact-value dichotomy from 20th century philosophers, particularly Hilary Putnam and Alasdair MacIntyre, who argue that evaluative judgments are inevitable, even in statements of fact, and should not be avoided. Their critiques may provide BCA practitioners with a means of escape from their dilemma. Section 4 discusses examples of how their critiques of the fact-value dichotomy have in fact been relevant in the history of BCA. Finally, Section 5 revisits the issue of distributional analysis in light of these insights.

2. Objectivity in economics: Separating values from facts

2.1. Wesley Clair Mitchell and the NBER

In 1920, Wesley Clair Mitchell (1874–1948) founded the NBER. Although today considered a “mainstream” economic organization, at the time the NBER represented something new. In the early 20th century, neoclassical economics was very deductive. Critics could persuasively accuse it of being mere armchair theorizing based on premises about human self-interest. In contrast, institutionalist economists like Mitchell claimed the mantle of science: they gathered data with which to generate and test theories. In Mitchell's words, they were replacing the “man of hunches” with the “man of facts” (quoted in Smith, 1994, p. 29).

Mitchell spent most of his career at Columbia, but he had studied economics at the University of Chicago, where he was influenced by Thorsten Veblen and John Dewey.⁶ Veblen was an institutional economist, from whom Mitchell learned to appreciate an empirical and historically grounded approach to economics. Dewey was a pragmatist philosopher who questioned the sharp dichotomy between science and ethics. On the empirical side, he argued scientific theories should be judged, not so much by how well they conform to reality, as by how useful they are for predictions and applications. He further argued that means and ends are in dialogue: just as we evaluate specific actions based on how well they advance our ends (or objectives), we revise our understanding of our ends based on

⁵ Hume (1739).

⁶ On Mitchell's background and philosophy of social science, see Burns (1952), Antler (1987), Hirsch (1988), and Biddle (1998).

our experience with attempts to arrive at them. Dewey held a high view of democracy as well as a progressive vision for science, which he saw as distinct but ideally complementary endeavors (Festenstein, 2018). The extent to which Mitchell accepted Dewey's means-end continuum is unclear (Biddle, 1998). But Mitchell did adopt Dewey's approach to empirical work as involving judgments about what is useful. He also esteemed democracy as distinct from science but in dialogue with it.

Mitchell's applied research matured during WWI, when he joined the War Industries Board. The board was trying to understand how to mobilize production for the war effort, but as its chief Bernard Baruch put it, "The greatest deterrent to effective action was the lack of facts" (quoted in Smith, 1994, p. 91). It is just hard to know how many resources you can mobilize if you do not even know the resource base. This experience shaped Mitchell's understanding of how applied work could contribute to public policy. Two years after the war, Mitchell founded the NBER, with economists such as Arthur Burns and Simon Kuznets also playing important roles in its early years.

In Mitchell's view and others', the democratic process makes normative decisions; social science finds the facts to inform that process. In that way, social science could be of realistic, practical use to business and government. To instantiate this philosophy, Mitchell established three rules for the NBER at its founding. First, it should only conduct research with public relevance. Accordingly, the NBER focused on such issues as business cycles, measuring inflation, and measuring national income. For example, its first report was titled, "Income in the USA: Its Amount and Distribution." (As is clear from the title, distributional issues were of concern from the outset of national income accounting.) Second, the NBER would present scientific facts free of all bias and propaganda. Accordingly, it tended to publish enormous volumes with numerous tables and charts. Third, it would make no policy recommendations, presenting only the facts to the true policymakers. This rule is still in force today.⁷

In fact, Mitchell was a progressive who wanted a more equitable income distribution. But he did not feel that it was his role, *qua* scientist, to advocate for particular policies. Rather, he believed his research could best serve policymaking by allowing it to speak for itself. Economic science should inform public decisions, but it should always be mindful of its place.

2.2. Lionel Robbins and the nature of economic science

A second episode surrounds Lionel Robbins (1898–1984), the LSE economist who shaped economics through his proposed redefinition of the field. In his *Essay on the Nature and Significance of Economic Science* (1935), Robbins defined it as a "science which studies human behavior as a relationship between ends and scarce means which have alternative uses." In other words, economics is a way of thinking about the world in terms of opportunity costs (p. 16). Robbins's *analytical* definition stood in contrast to the prevailing *classificatory* definitions at the time, which demarcated a set of economic topics. For example, according to Marshall, economics is the study of the "ordinary business of life," which he reduced to material rewards (Marshall, 1946, p. 14). Edwin Cannan, Marshall's counterpart at LSE and Robbins's predecessor, similarly equated economics with the study of wealth, which in turn relates to "material welfare" (Cannan, 1922, pp. 1–3).

⁷ <https://www.nber.org/about-nber/history>.

Robbin's definition is easy to take for granted today, but arguably it was one of the most important moves in the history of 20th century economics, shaping the profession in ways that were constraining in some respects but freeing in others. It was constraining in the sense that it restricted the range of economically acceptable logic. In this respect, it was hand-in-glove with the development of a post-war neoclassical orthodoxy. Under a definition of economics as material welfare, Mitchell clearly qualified as an economist and so, even, did Marx, as they manifestly were looking at material welfare. Under Robbins's definition, they are written out, with "Marxist economics" now almost an oxymoron.

On the other hand, the definition was freeing as it allowed a whole range of new topics to be explored using the economic way of thinking. After all, opportunity costs are everywhere. Robbins himself recognized this, bringing up the possibility of the economic study of such varied topics as religious practices or prostitution. Indeed, he had come to the idea in part by reflecting on his own war-time service and his dilemma at finding himself in the odd position of lecturing on the "economics of war." War and military strategy were not topics that fit well in the old definition of economics, but, again, they clearly involved using scarce resources. Thus, it is no surprise that Robbin's definition found mainstream acceptance in the 1960s, the same decade when Gary Becker was developing an economics of discrimination or crime, or a host of other things where there was not an economics of *that* before (Backhouse & Medema, 2009a, b).⁸

In his methodological essay, Robbins tried to walk a very fine line (Hands, 2009). On one hand, he wanted to avoid the discredited and, by 1930, frankly embarrassing psychological theories on which 19th century utilitarians and marginalists had built much of neoclassical theory. Those theorists thought they were following the science, but the science kept changing. Notably, Bentham had argued that human beings are motivated by hedonistic pain and pleasure tied to sense experience, but it is hard to square the content of that theory with even casual experience of human behavior. Jevons and Edgeworth grounded the more specific theory of diminishing marginal utility in 19th century psychophysics, especially the so-called Weber-Fechner law, according to which people display diminishing sensitivity to stimuli. For example, as we pile increasing weight into a pack, the additional weight required for a person to perceive the difference is proportionate to the weight already in the pack. According to Jevons and Edgeworth, hedonistic pleasure thus also diminishes with the "input" of consumption of a given good. Again, by the 20th century, plenty of counter-examples had been documented. Thus, economics was coming under attack for its shaky psychological foundations. Even worse, the very testability of economic theory, or at least its psychological foundations, was now being questioned, as there seemed to be no way to observe mental states and so no way to test utility theory. Thus, by the logic of positivism, talk of utility was nonsensical.

One possible response to these attacks was to accept methodological behaviorism, which was popular in psychology at the time. Behaviorism was an effort to look only at the outwardly observable behavior of agents without referring to their inscrutable states of mind or motives. This, in fact, was the road that Samuelson was contemplating around the same time, when introducing his axioms of choice behavior as a graduate student. (What we now know as the weak axiom of revealed preference was not originally called that by Samuelson.

⁸ On Robbins's intellectual journey to this definition, see Howson (2004). In previous work (Banzhaf, 2019, 2023), I have discussed how John Krutilla exploited these developments to create a new field of environmental economics, where, previously, it had been a question of economics versus the environment.

In fact, he was trying to rid economics of all reference to utility or preference.) But Robbins did not want to go that far, because doing so would rule out valuation and even the adjective *choice* modifying the behavior (pp. 87–88). Economic behavior would be indistinguishable from the theory of operant conditioning introduced by the great behaviorist B.F. Skinner or even from Pavlovian conditioning (Pavlov’s dogs).⁹

Instead, Robbins based his economic science on a particular kind of observation, namely, introspection or “inner experience.” Robbins said that we know from our inner experience that we can, and do, prioritize our alternatives. This “indisputable fact,” together with the facts that means are scarce and can be put to alternative ends, make up the foundation of Robbins’s economic science (pp. 12–13). These facts allowed him to keep the psychology of choice and valuation, without needing hedonism on the one side or behaviorism on the other. Said Robbins, “we can judge whether different possible experiences are of equivalent or greater or less importance to us. From this elementary fact of experience we can derive the idea of the substitutability of different goods, of the demand for one good in terms of another, of an equilibrium distribution of goods between different uses, of equilibrium of exchange and of the formation of prices” (p. 75).

Robbins’s push to move economics from a utilitarian foundation built on experienced pleasure and pain to a theory of choice was a key step in the second wave of the ordinal revolution of the 1930s, which he advanced along with his LSE colleagues R.G.D. Allen and John Hicks. With this ordinalist move, economics no longer needed to rely on a doctrine of hedonism or any other discredited psychology. It did not even need diminishing marginal utility (Robbins, 1935, p.92).

Robbins was very clear that people do have values, and that economics takes those values as facts. Adopting Max Weber’s thesis, Robbins conceded that in that sense economics is *value-laden*, but in a more important sense economics itself is *value-free* because it “is entirely neutral between ends” (pp. 83–91). “The ends may be noble or they may be base. They may be ‘material’ or ‘immaterial’ . . . But if the attainment of one set of ends involves the sacrifice of others, then it has an economic aspect” (pp. 24–25). Economics takes the ends or values as given and studies them in relation to scarce means.¹⁰

Robbins’s methodological positivism had important implications for what economic science has to say about economic policy. For example, does economic theory, per se, provide a rationale for redistribution? His mentor Cannan had said it did, introducing an argument that had recently been developed more thoroughly by A.C. Pigou in *The Economics of Welfare* (Pigou, 1932). To wit, because of diminishing marginal utility – and under the additional premise that any two individuals could achieve the same level of utility if given the same money incomes and prices – we can maximize total utility in the society through a more equitable distribution of money. After all, the rich person would lose money at a low level of marginal utility and the poor would gain money at a high level.

However, Robbins’s analysis forced the conclusion that economic theory had no such significance, for three reasons. First, we do not really need to assume anything about diminishing marginal utility, which is a cardinal concept. We only need the premise, verified by introspection that people rank alternatives in some ordinal way (pp. 138–139). Second,

⁹ For discussion of these points, see Hands (2009). For discussions of Samuelson’s moves and commitments (or lack thereof), see Hands (2014) and Moscati (2019).

¹⁰ For a summary of Weber’s theory of social sciences and objectivity, see Reiss and Sprenger (2020). Similar ideas are often articulated by economists (e.g., Gul & Pesendorfer, 2008).

interpersonal comparisons of utility are *unscientific*, because they cannot be verified either by external observations of behavior or by introspection (which can only inform us of our own minds). Such a comparison as one person's marginal utility of money to another's "necessarily falls outside the scope of any positive science... It involves an element of conventional valuation. Hence it is essentially normative. It has no place in pure science" (p. 139).

Third, even if we allow interpersonal comparisons of utility and concede that people have roughly equal capacities for enjoyment and diminishing marginal utility, we still could not conclude that we ought to redistribute money from the rich to the poor, because that would be basing a normative claim on facts, and, as Hume had said, we can never infer ought from is (pp. 142–143, 148–149).

In summary, Robbins's economic method differed completely from Mitchell's, the former emphasizing introspection and deductive theoretical models and the latter emphasizing the gathering of facts and inductive hypothesis formation. Yet both critiqued those aspects of utility theory that they deemed unscientific, and in the process severely limited policy recommendations. Both thought that economists, as scientists, can describe the distribution of resources and can forecast distributional changes under different scenarios. They can take as a datum what people think, ethically, about different distributions. But they cannot make policy recommendations, nor even employ a weighting system that suggests one distribution outperforms another.

3. Critiques of scientism

3.1. Risking irrelevance

The tight strictures that Mitchell and Robbins put on policy recommendations, in an effort to stay on the right side of the fact-value dichotomy, were not acceptable to everybody. They were attacked by others, on at least two fronts.

The first line of attack was that they risk making social science irrelevant. With respect to Mitchell's proposals, his junior colleague Robert Lynd, the Columbia sociologists famous for his study of *Middletown*, criticized Mitchell in his 1938 Stafford Little lectures at Princeton, published as *Knowledge for What? The Place of the Social Sciences in American Culture* (Lynd, 1939).¹¹ Lynd critiqued Mitchell and others who understood social science as pure technique. He argued that, if they confine themselves to studying the facts as they are, social scientists will accept the status quo instead of adopting a critical stance. For example, they will come to accept consumer culture and the existence of business cycles as inevitable. In contrast, with a more critical stance, they have the freedom to imagine how society might be if rightly ordered – that is, if it satisfied the true ends of consumption instead of stimulating false desires with advertising in order to feed production, as if production were an end in itself.

Lynd also wanted to demonstrate how purposively conceived, but still empirical, social science could achieve valued norms. For example, in his earlier work in *Middletown*, he argued that, whereas foolish, invidious comparisons to sophisticated urban life made people unhappy, contentment with a simpler communal life could lead to greater fulfillment.

¹¹ For a detailed discussion of Mitchell and Lynd, see Smith (1994).

Finally, Lynd wanted to fulfill John Dewey's plea to understand means and ends as part of a continuum. He thought that, yes, science could inform people about how to achieve given ends, but he also thought it could shape our ends.

Ultimately, Lynd's purpose-driven work (or, at least, the progressive purposes he particularly advanced) proved controversial and cost him his relationships with foundations. His career in academia continued to thrive, but his stance cost him the potential to serve in government or other spheres close to policymaking. This, of course, was precisely what Mitchell had feared, with good reason.¹²

Robbins met with similar criticisms about potential irrelevance but with very different solutions. Roy Harrod (1938) used his address to Section F of the British Academy, published in the *Economic Journal*, to reflect on Robbins's treatise. He pointed out that all economic policies have distributional effects. For example, movements toward free trade, such as the abolition of the corn laws that the classical economists had worked so hard to achieve, result in winners and losers. Such moves, he said, clearly result in a net increase in national income, but, if it is impossible to weigh the gains of one group against the losses of another, then it is impossible to state whether such reforms result in economic gains to society. Harrod could find no flaw in Robbins's logic. But he lugubriously lamented it and suggested there must be a role for a second branch of economics, which examines optimality, in addition to a first branch that looks at price formation.

In a remarkable sign of the efficiency of the exchange of ideas in the age of snail mail, Robbins, Kaldor, and Hicks responded with successive comments to one another in each issue of the *EJ*. First, in a model of handwringing, Robbins (1938) confessed that he felt Harrod's pain but had to go uncompromisingly where logic led him. In response, Nicholas Kaldor (1939) and John Hicks (1939) proposed what is now known as the potential Pareto test, or potential compensation test. Agreeing partially with Robbins, Kaldor conceded outright the impossibility of interpersonal comparisons of utility. But he argued both Robbins and Harrod had too quickly leapt from there to pessimistic conclusions about whether "economics as a science can say anything by way of prescription" (p. 549). In fact, he said, when "a certain policy leads to an increase in ... aggregate real income, the economist's case for the policy is quite unaffected by the question of the comparability of individual satisfactions; since in all such cases it is possible to make everybody better off than before..." (p. 550). Thus, according to Kaldor (and Hicks), the compensation test not only allows economists to say more, it allows economists to have it both ways, to make normative recommendations in the name of value-free science!

This history is tremendously important for understanding BCA's position as it moves more seriously into considering distributional issues. Crucially, the potential Pareto test was *not* developed merely to sidestep the need for economists to make an ethical value judgment about the distribution of income, while kicking the can down the road and leaving it to, say, the "distribution branch" of government (Musgrave) to address the question. That may have satisfied somebody like Mitchell. But Robbins had taken things a step further. It was not just that it is not *economists'* place to make such judgments. No.

¹² According to Biddle (1998), Mitchell's view was more subtle than he often led on, and he shared Lynd's recognition of a means-ends continuum, although he emphasized the distinction. He also was more cautious about overstepping.

By the lights of logical positivism, only statements that are observationally verifiable (or statements of pure logic) are intelligible, so any normative discussion about the relative merits of different distributional outcomes involves nonsense. Thus, it is impossible for *anybody* to make any such judgments rationally. Furthermore, since real-world policies always involve winners and losers, it is thus impossible for anybody – economist or policymaker – to rationally judge whether a policy results in a social improvement. The potential Pareto test was devised as an escape from this radical conclusion. It provided a way to tie statements about social improvements to facts, without making any ethical statements about distribution.

3.2. Fact-value entanglement

Meanwhile, even as economists were trying to honor this sharp fact/value dichotomy, philosophers in the postwar period were beginning to attack it and, along with it, Hume's no-ought-from-is dictum. Hilary Putnam (1926–2016), for example, another admirer of Dewey and a philosopher who paid a good deal of attention to economics, argued that science inherently *entangles* facts and values (Putnam, 2002; Putnam & Walsh, 2012).¹³ For one thing, scientists must conduct their work, including fact-finding work, according to their *epistemic values*, most overarchingly holding truth as a valued end, but also holding such subordinate values as parsimony, generality, coherence, elegance, predesigned research plans, and so forth. Additionally, they must make *evaluative judgments* when describing the world as it is, that is, when stating facts. To put a recent twist on one of Putnam's examples, take the descriptive statement, "the Ukrainian people are being very courageous." Stating such a fact inherently involves an evaluation of behavior as well as an evaluation of what it means to have courage, as opposed to, say, either foolhardiness on one side or cowardice on the other.¹⁴ Moving from such everyday descriptions to policy analysis, regulators must evaluate concepts such as "an adequate margin of safety;" economists must evaluate behavior as rational or irrational.¹⁵

Of course, we cannot just conflate fact and value either. Putnam still made a *distinction* between them, just not a sharp *dichotomy*. Some descriptions are more evaluative than others, and some of those evaluations have more ethical content than others. "The Ukrainians are courageous" is a thick description that entangles a good deal of ethical content about the virtue of courage in its description. It is simply impossible to make such a description without entangling values. The same can sometimes be true in economics. For example, descriptions of economic "development" inevitably involve evaluation (Nussbaum & Sen, 1989). Too, as Amartya Sen pointed out early in his career, some value judgments are tied up in factual content more than others. For example, the judgment that this year's increase in GDP is good depends on a number of facts, not only about the

¹³ See also Scarantino (2009) and Su and Colander (2013) for overviews intended for economists as well as (in the latter) a discussion of Putnam and Walsh's long-running methodological debate with Partha Dasgupta.

¹⁴ A classic question in philosophy (e.g., Plato, 1924; Aristotle, 1934, pp. 6–9).

¹⁵ McGartland (2021) discusses the problem of what he calls "science policy," and argues we should try to avoid such evaluations. Sugden (2009) notes that Robbins also tried as hard as he could to avoid any such evaluation, relying on introspection and questioning both the need or the possibility of experimental or econometric tests. On this point, he contrasts Robbins with Pareto and Wicksteed.

aggregate index itself but also related facts such as changes in the distribution of income (Sen, 1967).¹⁶

A second philosophical critique of the fact/value dichotomy has come from Alasdair MacIntyre (b. 1929), who takes a neo-Aristotelian teleological approach to critiquing Hume's no-ought-from-is dictum. A classic example is that from the factual description, "he is a sea captain" we might well conclude, "he ought to do whatever a sea captain ought to do." This is a teleological argument, because the ought is entangled with the end (*telos*) or purpose of seafaring. This example lacks specific content, but it at least makes the grammatical point that one can conclude an ought from an is. In *After Virtue* (1984), MacIntyre developed the argument with the aid of additional examples:

From such factual premises as 'This watch is grossly inadequate and irregular in timekeeping' and 'This watch is too heavy to carry about comfortably,' the evaluative conclusion validly follows that 'This is a bad watch.' From such factual premises as 'He get a better yield per acre for his crop than any other farmer in the district,' 'He has the most effective programme of soil renewal yet known' and 'His dairy herd wins all the best prizes at the agricultural shows,' the evaluative conclusion validly follows that 'He is a good farmer' (MacIntyre, 1984, pp. 57–58).

These arguments are valid because a watch and a farmer are functional concepts, defined by the functions they are expected to serve, so a farmer cannot be defined independently of a good farmer.

How does this line of argument apply to economics? Well, first, if we say that individuals are utility maximizers, then it follows that individuals ought to maximize their utility. From there, it is a short step to concluding that public policy ought to be directed to allowing them to do that, by increasing income, relaxing other constraints, providing public goods, and so forth. This logic is clearly implicit in Bentham's brand of utilitarianism, which seeks to find "the greatest good for the greatest number." It is even implicit in the strict Pareto test, which entails maximizing the good for one person, holding others' constant. Thus, according to MacIntyre, it is not so easy to wall off economic descriptions from their normative implications.

Similar critiques could be made of other schools of economics as well. For example, the classical economists routinely thought in terms of social groups or classes more than individuals, but these models implied their own respective oughts. For example, François Quesnay and the French physiocrats thought in terms of three groups: the land-owning class of proprietors, the productive class of farmers, and the unproductive class of artisans and merchants. The physiocratic scheme implied it was the duty of the proprietors to steer resources toward the productive class. Marx is another obvious example of somebody who thought in terms of groups, and the relationship between labor and the means of production. He also thought teleologically, arguing that History is governed by material forces which inevitably lead from feudalism to capitalism to socialism. The implication, of course, is that one ought to be on the right side of History.

¹⁶ Sen further argued that, in assuming the opposite, that is, in assuming that people *do not* revise ethical judgments on the basis of new facts, Robbins was himself adopting a premise that was not falsifiable. See Putnam (2002, Ch. 4) and Walsh (2003) for additional discussion of Sen's critique of Robbins (see Sen, 1987 for a broader statement about his view on the relationship between economics and ethics).

I am not saying that any of these are particularly good models. My point is only that neoclassical economics is not the only school that faces is-ought dilemmas. They are inherent to economics because economics always involves thick descriptions with normative implications.

4. Fact/value entanglement and BCA

4.1. What “counts” as an economic value? The case of outdoor recreation

Such fact/value entanglement can affect BCA in at least two ways. First, it affects what counts as a benefit or cost. This issue has been present from the very first attempts to systematize BCA, when American practitioners created the so-called “Green Book,” a 1950 handbook on BCA for water projects.¹⁷ The practitioners had found that, across the wide array of agencies working on such projects, there was an equally wide array of inconsistent practices. Such inconsistencies are unsettling, as they undermined the authority that BCA needs to have if it is to settle disputes (Porter Theodore, 1995). As John Maurice Clark and others put it in a consultant’s report: “Democracy has to rely on technicians in matters inscrutable to the non-specialist, but preferably where the specialist is following a well-authenticated technique. In this case, the disagreements among the specialists are evidence that they do *not* possess an authenticated technique...” (Clark *et al.*, 1952, p. 11). The Green Book was an effort to settle the disputes and authenticate BCA techniques.

How did the Green Book economists think about benefits? In terms of the general theory, they thought in terms of national development and national income. Accordingly, they were thinking about incomes to firms and rents to factors of production. In other words, they were thinking of *revenue* rectangles, not something like consumer surplus. Moving to specifics, they began by identifying the key types of benefits of water projects, including reclamation, hydroelectric power, flood control, navigation, municipal water supplies, and outdoor recreation.

Quantitatively, recreation was by no means the most important, but it received a great deal of attention. This attention can be attributed to five factors. First, and most importantly, recreation was becoming increasingly important, as participation accelerated following the War (Clawson, 1958). Second, frankly, including recreation would pump up the benefit-cost ratios, and there was a lot of pressure to do that. A third factor was the role of recreation in cost-sharing formulae. Since the Reclamation Act of 1902, reclamation projects required reimbursement from farmers. But costs that were explicitly incurred for other objectives did not need to be recovered, nor did the full share of *joint costs* incurred for multiple objectives. So if some costs could be allocated to recreation, it would reduce the net costs to farmers. In this way, quantifying outdoor recreation benefits translated into farm aid – always popular to governments. Fourth, handling recreation created some frictions across agencies when their incentives were not aligned. Finally, because it is a nonmarket good, recreation was just one of the most difficult nuts to crack.

Each of these factors can be illustrated with a single episode, centered around the proposed Echo Park dam in Dinosaur National Monument, near the Colorado/Utah border. It was one of the most important episodes in the history of American environmentalism, and

¹⁷ This section is based on Banzhaf (2010, 2023).

a first big win for the movement. It was also the spur to serious thinking about the economics of outdoor recreation.

When first proposed by the Bureau of Reclamation before the war, the dam seemed like a common-sense use of public lands to obtain much-needed water resources in the west. The National Parks director, Newton Drury, was at first supportive. But after the War, popular attitudes were changing, and public opinion turned against the project. By 1955, it was dead. But, meanwhile, it was caught up in the interagency clashes inside the Department of the Interior. The whole project was problematic from the beginning: it was a Bureau of Reclamation idea, advancing its mission, but on NPS land and taxing its staff. Although the precise details of when and why are murky, by the late 1940s Drury had turned against the project and begun to fight it. This reversal won him few friends, and eventually, he was forced to resign in 1951.

Intriguingly, during these machinations, Drury exploited the uncertainties surrounding how to measure recreation benefits. In 1948, the NPS concluded that, in general, the values of park and recreation facilities cannot be measured quantitatively, so they would “no longer attempt to furnish such estimates.” Instead, only a “judgment value” would be made. Furthermore, for expediency, the NPS would simply use the blanket rule that recreation benefits were equal to those costs specifically incurred for recreation facilities. (After all, why else would they do it if the benefits were not at least that amount?) Moreover, no portion of joint costs would be allocated to recreation.

Reclamation leaders were apoplectic. As they pointed out, the expediency would dilute benefit-cost ratios, as a ratio of one was being averaged in. Additionally, by minimizing the cost share allocated to recreation, it made farmers and ranchers (i.e., Reclamation’s stakeholders) bear more of the burden. Furthermore, even if any joint costs subsequently were allocated to recreation, because benefits were set equal to direct (not joint) costs, the benefit-cost ratio for the recreation portion would then be less than one, thus creating a catch-22.

The Secretary of Interior clearly was displeased, and Drury eventually backed off. Thus chastened, the NPS reversed itself: from now on, it would assume that benefits were equal to 2× costs! This absurd formula clearly undermined the scientific credibility of BCA. One cynical – but entirely plausible – explanation of Drury’s move is that he made it so as to appear to be cooperative, while adopting a plan that he must have known the economists would shoot down, thus doing his dirty work for him. Indeed, the Green Book economists and others throughout the bureaucracy did just that, sending the valuation methodology back to the drawing board.

Meanwhile, for the economists just trying to keep their heads down and do their jobs, the essential problem of valuing recreation remained: prices are not observed. In 1948, to help it find a way out of its dilemma, the NPS elicited the opinions of 10 prestigious experts. Of the 10, only Harold Hotelling held out hope for finding a way to estimate recreation benefits. His response is the source for the modern Travel Cost Model of recreation demand, but at the time his minority recommendation was ignored by the NPS. The other nine experts were unanimous that it could not be done.

Based on their recommendations, Park economist Roy Prewitt concluded that it would be better not to measure recreation benefits:

Recreation is, first of all, an intangible – a service. It is not a standardized or homogeneous service; it varies with every individual and it cannot be considered separate and apart from the individual. It is of the mind and body, it cannot be stored or

transported, it is a psychic value and it cannot be measured in objective terms. Finally, the recreational values supplied by the National Park Service are not sold for a price under marketplace rules. (NPS, 1949, p. 12).

Prewitt concluded that “it might be better to forget the words ‘economic value of recreation’ and focus attention on the *expenditures induced* by recreation.” For, “It is in this area that an *objective* approach can be made...” (emphasis in original). The next year, NPS began gathering data on daily expenditures for recreation trips. In 1957, it officially adopted this “unit-day” approach.

This episode illustrates four important realities about the relationship between BCA and the wider policy and social environment. First, BCA is subject to pressures from the crudest forms of interest group politics. In this case, that pressure came from petty bureaucratic turf wars, but also from affected stakeholders like landowners who wanted Western development and farmers who wanted offsets to their Reclamation reimbursements. The Green Book economists, and BCA practitioners today, are right to resist these kinds of pressures, upholding their objectivity in that sense of the word. Second, though, BCA truly is entwined with political questions, in the best sense of the word “political.” That is, it is entwined with debates about how to govern the polis. Facts documented by BCA can inform those debates, while at the same time, the people’s concerns and values shape the analysis.

Third, how BCA practitioners define a concept like “economic benefits” has implications for which values get incorporated into our analyses, and consequently whose values are counted. Thus, they cannot describe benefits (or costs) as pure facts without also entangling values. At a time when “economics” meant material welfare, practitioners evaluated the immaterial enjoyment of wilderness as out of bounds. Too, at a time when “benefits” meant revenues, they preferred to quantify the expenditures induced by recreation to quantifying the consumer surplus from the free services afforded by such experiences.

Finally, it can be a fine line between striving to measure some object objectively and choosing the object to be measured based on the objectivity of the measurements. That is, objectivity is an epistemic value that is not necessarily neutral. In this case, the objectivity of the measurements shaped the decision to measure expenditures rather than benefits. For, as Prewitt put it, recreation “is of the mind and body. . . it is a psychic value and it cannot be measured in objective terms.” Only a subjective judgment value can be made. Measuring expenditures was preferable because “it is in this area that an *objective* approach can be made.” Similar issues have plagued other kinds of valuation questions over the years, such as the value of life.

4.2. Whose values count? BCA and distributional effects

These implications for *what* we value necessarily entail implications for *whose* values count, and to what degree. For example, in the seminal case of outdoor recreation, whether we count it or not determines the degree to which we count the values of recreationists. Whether we count health effects determines the degree to which we count the values of those whose health is jeopardized, and so forth.

This brings us back to the issue of interpersonal comparisons. The description of society as the sum of individuals entangles facts and values in particular ways, but the efforts by Robbins and Kaldor and others to disentangle them have led to some of the biggest dilemmas we face in BCA. Those dilemmas arise because, by trying to separate out distributional

effects, which require ethical assessments, from the facts of net benefits, they only buried them deeper. This can be seen in three ways.

First, as decision rules for ranking social states based on individual noncomparable ordinal preference relations, potential Pareto tests cannot escape Arrow's (1951) impossibility theorem. They are not rational choice orderings. In particular, as Tibor Scitovsky showed, they can have intransitive cycles even when each individual has well-ordered preferences, even cycles of two.¹⁸ That is, it is entirely possible that there could be two points in the policy space, A and B, such that a move from point A to point B, accompanied by a suitable redistribution of resources, would make everybody better off *and* that a move from B to A, accompanied by a different redistribution, would make everybody better off. These are known as Scitovsky reversals. These reversals arise because, fundamentally, it is impossible to aggregate noncomparable utility functions in a way to rationally order social states, which inevitably involve tradeoffs among individuals.

Second, *potential* compensations are not the same thing as *actual* compensations (Little, 1957; Sen, 1987). It is small consolation to the poor to be told that Policy X will impose great costs on them but will be adopted because it increases economic efficiency enough to potentially compensate them, though there are no plans to do so! Thus, when used as a decision rule, the potential Pareto test is effectively a normative assumption that distributional assumptions do not matter.

In fact, one could even use the logic of potential Pareto tests to block any proposed redistribution program or repeal existing ones. Any real-world taxation-and-redistribution policy has some deadweight loss, some excess burden, associated with it. Therefore, repealing it would increase economic efficiency, removing the deadweight loss and creating enough surplus to allow the winners to compensate the losers through lump-sum transfers. Only a truly lump-sum transfer could pass the test, but such policies do not exist. To address this concern, Harberger (1980) proposed replacing the lump-sum transfers of the standard potential Pareto tests with the least-cost, real-world transfer scheme, but even his proposal would still justify repealing a relatively inefficient transfer scheme on the grounds that it would be possible to find a better one, without actually doing the better one.

Third, as a matter of logic, even the *strict* Pareto test involves the normative evaluation that it is good to increase everybody's real income, a point that Hicks himself came to realize (1975). Thus, tests of potential Pareto improvements entail thick descriptions of the type discussed by Putnam. They entail judgments about the potential to actually make a transfer as well as judgments about what constitutes an improvement.

Another way of seeing this is to go back to Robbins's defense of economics as the science of *choice* behavior on the grounds that we can rely on the fact, verifiable from introspection, that people have preference orderings. He might just as well have concluded from introspection that people have preference orderings for *social states* including distribution. Without explicitly discussing it, Robbins seems to have made the evaluative judgment that these preferences do not count, which is hardly the scientific thing to do. In this sense, his overall argument for noncomparability is self-refuting.

Benefit-cost practitioners have tried to counter the undesirable ethical properties of the potential Pareto test with other moves that are hard to defend on economic grounds, in the

¹⁸ Scitovsky (1941), Samuelson (1950), and Gorman (1955). See Blackorby and Donaldson (1990) for an excellent overview.

hopes that two wrongs make a right. Consider for example our standard use of a homogeneous value of statistical life (VSL) for everybody, regardless of age or income. This hardly matches the empirical facts, where we find different groups willing to make different tradeoffs, because of differences in their ability to pay, their access to substitutes, or preferences.

Of course, it is understandable why we do this. Because we cannot (or do not) use distributional weights, and because the rich have higher VSLs than the poor, introducing heterogeneous VSLs would systematically tilt BCA in favor of policies that benefit the health of the rich over the poor. Needless to say, such procedures would raise serious ethical objections. Ignoring heterogeneity in VSLs seemingly solves this problem. However, it creates a new one. As I illustrate below, ignoring heterogeneity in VSLs can make a policy look like a Pareto improvement when in fact all groups, including the poor, are worse off under it! Thus, standard BCA decision rules can reduce welfare and fail to do the justice of honoring a group's right to set its own priorities. This dilemma turns out to be rooted in the premise that we cannot or should not use distributional weights. One solution is to use both distributional weights *and* heterogeneous VSLs.

Consider the following numerical example.¹⁹ Suppose the average VSL is \$6 m, but the VSL of the rich is \$8 m and the VSL of the poor, because of their lower income, is \$4 m. The poor cannot afford to pay as much money to reduce risks to their health and safety without foregoing other basic needs, while the rich can make such purchases while only foregoing luxuries.

Consider now two policies, Policies A and B, that save lives. Tables 1 and 2 show that both policies impose gross costs of \$1700 m on the rich but nothing on the poor. Policy A saves 100 statistical lives of the rich and 200 of the poor, for a total of 300. Policy B saves 200 lives of the rich and 50 lives of the poor, for a total of only 250. Because it saves more lives at the same cost, Policy A must look better when we use the average VSL of \$6 m, and indeed it does (\$100 m vs. -\$200 m, with net values in bold). If we use heterogeneous values, however, Policy A would generate \$-900 m in net benefits for the rich and only \$800 m in net benefits for the poor, for an aggregate loss of \$100 m. Policy B would generate \$-100 m in net benefits for the rich and \$200 m in net benefits for the poor, for a net gain of \$100 m in aggregate. Policy B has higher net benefits. Thus, using heterogeneous values, the efficiency criterion seemingly steers us to Policy B because it saves more rich lives. This would seem to imply that socially, we would trade 100 lives of the poor for 50 lives of the rich. Nothing could be less just or more reprehensible! It would seem the logic for using homogenous VSLs is clear.

Yet, in fact, *the supposed choice of Policy B does not follow from using heterogeneous VSLs per se, but only from doing so without distributional weights.* Giving greater weight to the net benefits of the poor would have steered us back to Policy A, which intuitively is the right choice. Why use heterogeneous VSLs if we are going to undo them with the distributional weights? The reason can be made clear with one more example.

Consider two different policies, C and D, also illustrated in Tables 3 and 4. Policy C costs \$650 m, with the costs falling \$250 m on the rich and \$400 m on the poor. Policy D costs \$700 m, with \$600 m falling on the rich and only \$100 m on the poor. Both policies save 150 lives, but Policy C saves 100 of the poor and 50 of the rich, while Policy D reverses the split, saving 100 of the rich and 50 of the poor. Using homogenous VSLs of \$6 m, we see that

¹⁹ These examples are variants of ones I previously used in Banzhaf (2011).

Table 1. Benefit-cost analyses for Policy A.

Group	Costs (m)	Lives saved	Benefits without heterogeneity (m)	Benefits with heterogeneity (m)	Net benefits without heterogeneity (m)	Net benefits with heterogeneity (m)
Rich	\$1700	100	\$600	\$800	−\$1100	−\$900
Poor	0	200	\$1200	\$800	\$1200	\$800
Total	\$1700	300	\$1800	\$1600	\$100	−\$100

Note: Benefits without heterogeneity in willingness to pay are based on a VSL of \$6 m; benefits with heterogeneity are based on a VSL of \$8 m for the rich and \$4 m for the poor. Net values in bold.

Table 2. Benefit-cost analyses for Policy B.

Group	Costs (m)	Lives saved	Benefits without heterogeneity (m)	Benefits with heterogeneity (m)	Net benefits without heterogeneity (m)	Net benefits with heterogeneity (m)
Rich	\$1700	200	\$1200	\$1600	−\$500	−\$100
Poor	0	50	\$300	\$200	\$300	\$200
Total	\$1700	250	\$1500	\$1800	− \$200	\$100

Note: Benefits without heterogeneity in willingness to pay are based on a VSL of \$6 m; benefits with heterogeneity are based on a VSL of \$8 m for the rich and \$4 m for the poor. Net values in bold.

Table 3. Benefit-cost analyses for Policy C.

Group	Costs (m)	Lives saved	Benefits without heterogeneity (m)	Benefits with heterogeneity (m)	Net benefits without heterogeneity (m)	Net benefits with heterogeneity (m)
Rich	\$250	50	\$300	\$400	\$50	\$150
Poor	\$400	100	\$600	\$400	\$200	\$0
Total	\$650	150	\$900	\$800	\$250	\$150

Note: Benefits without heterogeneity in willingness to pay are based on a VSL of \$6 m; benefits with heterogeneity are based on a VSL of \$8 m for the rich and \$4 m for the poor. Net values in bold.

Table 4. Benefit-cost analyses for Policy D.

Group	Costs (m)	Lives saved	Benefits without heterogeneity (m)	Benefits with heterogeneity (m)	Net benefits without heterogeneity (m)	Net benefits with heterogeneity (m)
Rich	\$600	100	\$600	\$800	\$0	\$200
Poor	\$100	50	\$300	\$200	\$200	\$100
Total	\$700	150	\$900	\$1000	\$200	\$300

Note: Benefits without heterogeneity in willingness to pay are based on a VSL of \$6 m; benefits with heterogeneity are based on a VSL of \$8 m for the rich and \$4 m for the poor. Net values in bold.

the aggregate net benefit of Policy C is \$250 m whereas for Policy D it is only \$200 m. Using the efficiency criterion alone, Policy C dominates. Moreover, it seems to raise no distributional red flags, as no group is worse off under C than D. Policy C looks more favorable, so using these criteria we would choose it over Policy D. But this is the wrong conclusion. When we consider the groups' true VSLs, we now see that *both groups* are better off under Policy D than Policy C. Under Policy D, the poor get \$100 m in net benefits versus zero under Policy C, while the rich get \$200 m versus \$150 m.

The problem with Policy C relative to D is that the additional 50 lives of the poor saved come at an incremental cost to the poor of \$300 m, while the group is only willing to pay \$200 m for those statistical lives. So it requires the poor to actually pay a cost they cannot afford: For them, more basic priorities (perhaps food and shelter) take precedence over the reduction in risks, whereas the rich can afford the cost.

The only way out of this dilemma and to make the seemingly right choice in both comparisons (A over B *and* D over C) is to consider *both* heterogeneity in willingness to pay *and* distributional objectives in the analysis.

5. Finding social weights

Unfortunately, incorporating distributional weights into BCA creates a new set of problems.²⁰ Where are the weights to come from? Generally, they have two potential bases (Del Campo *et al.*, 2023). One approach begins with first ethical principles and then adopts a social welfare function (or other kinds of weights) that reflect them. A second, more empirical approach looks to other social choices, such as income tax rates or income support programs, and fits a social welfare function that can best explain those choices.

The problem with the deductive approach, where practitioners impose a social welfare function on the analysis, is that it would take BCA practitioners across a fact-value Rubicon. It would impose their ethics upon the broader society, which they are not entitled to do. Consequently, it inevitably would invite criticism of BCA from any elements in society that object to the ethical principles imposed upon them. If the potential Pareto test can be criticized for naïvely separating facts and values, the social welfare approach can be criticized for ignoring the distinction altogether.

The problem with the empirical approach is that it lacks rational foundations. It looks to existing social choices as revealing society's ethical principles, but there is no reason to believe existing choices are particularly ethical. Moreover, it involves a certain circularity. It reforms BCA in such a way that new policies pass BCA tests if they are consistent with old policies. As time goes on, this process becomes increasingly self-reflective. Thus, rather than informing the decision-making process, it simply reflects that process back to itself. Policies will pass BCA tests if and only if they are the kinds of policies that policymakers make.

Interestingly, these issues were actually addressed in a previous attempt to reform BCA (Banzhaf, 2009, 2023). In 1969, the Water Resources Council (the successor to the Green Book committee) began to explore new Principles and Procedures for BCA of US water projects. They were to revise the existing standards, enshrined in 1962 in what was known as

²⁰ Adler (2012, 2016) reviews social welfare functions and their potential use in BCA. Fleurbaey (2009) gives a broader overview of measures of social welfare, including social welfare functions, Sen's (1992) capabilities approach, and other approaches.

Senate Document 97. But the process became bogged down in contentious debates about discount rates, nonmarket valuation – and distributional issues. Eventually, new guidance was finally adopted in 1973, which lasted 10 years until the Reagan-era guidance.

Two groups of economists were particularly active in the debates over the new BCA standards, one based at Harvard University's Water Program and another at Resources for the Future. The Harvard Water Program was an interdisciplinary group meeting since the 1950s. In 1962, they published a book titled *Design of Water Resource Systems* (Maass *et al.*, 1962). At that point, the team comprised Arthur Maass (a political scientist), Maynard Hufschmidt (public administration), Gordon Fair and Harold Thomas of engineering, and economists Robert Dorfman, Otto Eckstein, and, at first as an undergraduate student, Stephen Marglin.

Design advanced the idea of multi-objective BCA. This means more than the long-recognized fact that water systems have multiple *purposes*. Rather than aggregate them into a single number, Maass *et al.* (1962) wanted to leave as disaggregated such broad categories as national income, nonmarket goods such as environmental quality or lives saved, regional economic development, and individual-level inequality. As Dorfman put it, the relative values of human life, endangered species, and jobs in a particular way of life are “not questions of fact that might admit expert answers, but questions about social values and public preference, [which] only the elaborate and clumsy procedures of democratic decision-making can answer. Such answers are not data to be fed into decision-making processes but, rather, outputs of those processes” (Dorfman, 1997, p. 373).

How would this work? Maass proposed a four-step procedure in which, first, agency analysts represent the tradeoffs among various objectives across alternative projects or designs of a project; second, the executive proposes a program to the legislature; and, third, the legislature decides. In other words, analysts would merely make the production possibilities curve known, while elected representatives would make social choices. Maass further envisioned the process iterating back in a fourth step, with agency analysts improving their design of project alternatives based on past social choices. For a time, it looked like the new Principles and Procedures would entail a radical reform of BCA along these lines.

Meanwhile, two other economists, just beginning their careers, were writing their PhD dissertations on the question of efficiency-equity tradeoffs in the context of water resources policy: the lates Robert Haveman and Myrick Freeman. Haveman conducted retrospective assessments of how water resource dollars were actually spent by Congress. Using distributional weights based on marginal tax rates (as an indicator of Congresses' distributional tradeoffs), he found that more Army Corps projects passed a benefit-cost test than with unweighted BCA. Freeman conducted a similar exercise but used an explicit social welfare function of his own choosing. He found most Reclamation projects still failed.

Haveman and Freeman came together for a time at Resources for the Future in 1969 and found in one another kindred spirits. First, in their philosophical posture, they believed in reducing multiple objectives to a single social objective function, a function that aggregated the preferences of individuals over social states. Second, they believed that the political process was a poor way to meet these objectives, because of the familiar problems of the pork barrel and empire-building. While uniting them together, these themes served to separate them from the Harvard Water Program. Together with other RFF economists like Jack Knetsch, they engaged Maass *et al.* (1962) in a major debate over the shape and role of BCA.

Freeman, Haveman, and Knetsch argued that traditional BCA was the most useful tool for two reasons. First, it is the most *objective*. It captured the biggest benefits and costs that

deserved first consideration, whereas allowing multiple objectives in an analysis, some of them unpriced, would open up the process to political manipulation and abuse. Interestingly, Maass *et al.* (1962) argued the opposite, namely, that designing and evaluating projects on the basis of a single efficiency objective, when the legislature wanted to make its choice on the basis of multiple objectives, created the incentive for manipulation, to jiggle the numbers to account for qualitative factors. One might argue the use of homogenous VSLs is an example of this.

Second, the RFF economists argued that looking to Congress for weights, albeit an interesting exercise descriptively, could not provide a normative signal about what the weights *should* be, which would be crucial for improving policy. In fact, it was circular, and could always justify Congressional action. That is, it would make BCA a descriptive exercise, instead of a normative one. It could never advise political decision-makers nor evaluate their decisions based on an objective criterion, except maybe to identify strictly dominated projects.

Of course, this is true. It is very hard indeed to evaluate normative decisions on objective criterion.

In hindsight, there were big issues at play here: nothing less than the fate of two competing visions of liberalism and the role of the “expert” within them. The RFF economists emphasized the importance of *consumer sovereignty*, according to which benefits (including the satisfaction of preferences for more equity) accrued to individuals. From this perspective, the task of BCA is to make inferences from individuals’ behavior, and on that basis to give advice to policymakers. It can also be a basis to judge the performance of the political decision-makers, who can be too easily manipulated by special interests into making policies contrary to the commonweal. In contrast, the Harvard team emphasized the importance of collective choice and *political sovereignty*, in which elected officials represented the will of the people. From this perspective, the task of BCA is to observe policymakers, infer their needs, and provide them with the information that can facilitate best their decision-making.

6. Conclusions

From this intellectual history, we can draw three useful lessons as distributional issues become increasingly import in BCA. First, it is impossible to conduct BCA objectively in the sense of making it value-free. The concepts of “benefits” and “costs” inherently entangle facts and values. Thus, describing distributional effects does not cross some particular bright line. Nor does ignoring distributional issues make BCA any more objective. It simply embeds different implicit values, just as ignoring recreational benefits in the name of objectivity did some 75 years ago.

Second, BCA is used for social decision-making, so it should provide the kind of information that society considers relevant. As distributional issues manifestly are relevant to much of society, if we BCA practitioners *do not* consider distributional issues, we are going to be irrelevant. Indeed, this may have been one reason all along that we have not gotten the deference that we sometimes would have like, as reasonable people will ignore us if we are not addressing their concerns.

Third, BCA is inherently threatened by partisan politics and interest-group manipulation. Such manipulation becomes especially salient any time BCA ventures into new areas, without “a well-authenticated technique.” This was true in the early days of interagency

BCA, with the controversies that led to the Green Book; it was true when introducing nonmarket valuation in the evaluation of outdoor recreation; and, we might add, it continues to be true for quantifying health and mortality (McGarland, 2021). Likewise, incorporating distributional issues, as much as it is warranted on its own terms, will increase the risk of such controversy.

Thus, we BCA practitioners will have to tread carefully, avoiding irrelevance on one hand and the appearance of political bias on the other. As emphasized both by Mitchell and by members of the Harvard Water Program, we should think carefully about our place in policymaking. We provide the best service by being as transparent as possible about the ways values influence BCA reasoning, without arrogating political decisions. But if we pick one “preferred” social welfare function to introduce into BCA, that is exactly what we would be doing. Instead, we can focus on the descriptive, showing the tradeoffs among groups. Or, especially at first, we can give weighted BCA results under a range of possible weights, including the traditional potential compensation test, documenting sensitivity. In a democracy, important social tradeoffs need to be made by democratic authorities, not academic experts. In summary, BCA is a tool, not a rule.

Acknowledgments. This article began as a plenary session at the 2022 Meetings of the Society for Benefit-Cost Analysis. I thank G. Blomquist, C. Cecot, M. Cropper, L. Robinson, and other participants for their comments.

Competing interest. The author declares none.

References

- Acland, Daniel J., and David H. Greenberg. 2023. “Distributional Weighting and Welfare/Equity Tradeoffs: A New Approach.” *Journal of Benefit-Cost Analysis*, 14(1): 68–92.
- Adler, Matthew D. 2012. *Well Being and Fair Distribution: Beyond Cost-Benefit Analysis*. New York: Oxford University Press.
- Adler, Matthew D. 2016. “Benefit–Cost Analysis and Distributional Weights: An Overview.” *Review of Environmental Economics and Policy*, 10(2): 264–285.
- Antler, Joyce. 1987. *Lucy Sprague Mitchell: The Making of a Modern Woman*. New Haven, CT: Yale University Press.
- Aristotle. 1934. *Nicomachean Ethics, v. XIX of Aristotle*. Translated by H. Rackham. Cambridge, MA: Harvard University Press.
- Arrow, Kenneth J. 1951. *Social Choice and Individual Values*. New York: Wiley.
- Astrid, Matthey, and Björn Bünger. 2019. Methodological Convention 3.0 for the Assessment of Environmental Costs. Technical Report. German Environmental Agency. Available at www.umweltbundesamt.de/sites/default/files/medien/1410/publikationen/2019-02-11_methodenkonvention-3-0_en_kostensatze_korr.pdf.
- Backhouse, Roger E., and Steve G. Medema. 2009a. “Defining Economics: The Long Road to Acceptance of the Robbins Definition.” *Economica*, 76: 805–820.
- Backhouse, Roger E., and Steven G. Medema. 2009b. “Retrospectives: On the Definition of Economics.” *Journal of Economic Perspectives*, 23(1): 221–233.
- Banzhaf, H. Spencer. 2009. “Objective or Multi-Objective? Two Historically Competing Visions for Benefit-Cost Analysis.” *Land Economics*, 85(1): 3–23.
- Banzhaf, H. Spencer. 2010. “Consumer Surplus with Apology: A Historical Perspective on Nonmarket Valuation and Recreation Demand.” *Annual Review of Resource Economics*, 2(1): 183–207.
- Banzhaf, H. Spencer. 2011. “Regulatory Impact Analyses of Environmental Justice Effects.” *Journal of Land Use and Environmental Law*, 27(1): 1–30.
- Banzhaf, H. Spencer. 2019. “The Environmental Turn in Natural Resource Economics: John Krutilla and ‘Conservation Reconsidered.’” *Journal of the History of Economic Thought*, 41(1): 27–46.
- Banzhaf, H. Spencer. 2023. *Pricing the Priceless: A History of Environmental Economics*. Cambridge: Cambridge University Press, forthcoming.

- Biddle, Jeff. 1998. "Social Science and the Making of Social Policy: Wesley Mitchell's Vision." In Rutherford, Malcolm (Eds.) *The Economic Mind in America: 55–91*. London: Routledge.
- Blackorby, Charles, and David Donaldson. 1990. "A Review Article: The Case Against the Use of the Sum of Compensating Variations in Cost-Benefit Analysis." *Canadian Journal of Economics*, 23(3): 471–494.
- Burns, Arthur F. (Ed.) 1952. *Wesley Clair Mitchell: The Economic Scientist*. New York: NBER.
- Cannan, Edwin. 1922. *Wealth: A Brief Explanation of the Causes of Economic Welfare*, 2nd Ed. London: P.S. King & Son.
- Cecot, Caroline. 2023. "An Equity Blindspot: The Incidence of Regulatory Costs." *Journal of Benefit-Cost Analysis*, 14(1): 35–43.
- Clark, John M., Eugene L. Grant, and Maurice M. Kelso. 1952. *Report of Panel of Consultants on Secondary or Indirect Benefits of Water Use Projects*. Submitted to the Bureau of Reclamation.
- Clawson, Marion. 1958. *Statistics on Outdoor Recreation*. Washington, DC: Resources for the Future.
- Del Campo, Stelio, David Anthoff, and Ulrike Kornek. 2023. "Inequality Aversion for Climate Policy." *Review of Environmental Economics and Policy*, forthcoming.
- Dorfman, Robert. 1997. "Introduction." In *Economic Theory and Public Decisions*. Cheltenham: Edward Elgar.
- Elliott, Kevin C., and David B. Resnik. 2014. "Science, Policy, and the Transparency of Values." *Environmental Health Perspectives*, 122(7): 647–650.
- Festenstein, Matthew. 2018. "Dewey's Political Philosophy." In Zalta, Edward N. (Ed.) *The Stanford Encyclopedia of Philosophy*. Stanford, CA: Stanford University. Available at <https://plato.stanford.edu/archives/fall2018/entries/dewey-political/>.
- Fleurbay, Marc. 2009. "Beyond GDP: The Quest for a Measure of Social Welfare." *Journal of Economic Literature*, 47(4): 1029–1075.
- Gorman, William M. 1955. "The Intransitivity of Certain Criteria Used in Welfare Economics." *Oxford Economic Papers* 7(1): 25–34.
- Gul, Faruk, and Wolfgang Pesendorfer. 2008. "The Case for Mindless Economics." In Caplin, Andrew, and Andrew Schotter (Eds.) *The Foundations of Positive and Normative Economics: A Handbook*: 3–39. New York, NY: Oxford University Press.
- Hands, D. Wade. 2009. "Effective Tension in Robbins' Economic Methodology." *Economica*, 76: 831–844.
- Hands, D. Wade. 2014. "Paul Samuelson and Revealed Preference Theory." *History of Political Economy*, 46(1): 85–116.
- Harberger, Arnold C. 1980. "On the Use of Distributional Weights in Social Cost-Benefit Analysis: Reply to Layard and Squire." *Journal of Political Economy*, 88(5): 1050–1052.
- Harrod, Roy F. 1938. "Scope and Method of Economics." *Economic Journal*, 48: 383–412.
- Hicks, J. R. 1939. "Foundations of Welfare Economics." *Economic Journal*, 49: 696–712.
- Hicks, John R. 1975. "The Scope and Status of Welfare Economics." *Oxford Economic Papers*, 27(3): 307–326.
- Hirsch, Abraham. 1988. "What is an Empiricist? Wesley Clair Mitchell in Broader Perspective." *H.E.S. Bulletin*, 10(1): 1–12.
- Howson, Susan. 2004. "The Origins of Lionel Robbins's Essay on the Nature and Significance of Economic Science." *History of Political Economy*, 36(3): 413–443.
- Hume, David. 1739. *A Treatise of Human Nature: Being An Attempt to Introduce the Experimental Method of Reasoning into Moral Subjects*. Available at <https://oll.libertyfund.org/title/bigge-a-treatise-of-human-nature>.
- Kaldor, N. 1939. "Welfare Propositions of Economics and Interpersonal Comparisons of Utility." *The Economic Journal*, 49: 549–552.
- Kniesner, Thomas J., and W. Kip Viscusi. 2023. "Promoting Equity through Equitable Risk Tradeoffs." *Journal of Benefit-Cost Analysis*, 14(1): 8–34.
- Lakdawalla, Darius N., and Charles E. Phelps. 2023. "The Generalized Risk-Adjusted Cost-Effectiveness (GRACE) Model for Measuring the Value of Gains in Health: An Exact Formulation." *Journal of Benefit-Cost Analysis*, 14(1): 44–67.
- Little, I. M. D. 1957. *A Critique of Welfare Economics*, 2nd Edn. Oxford: Oxford University Press.
- Lynd, Robert S. 1939. *Knowledge for What? The Place for Social Science in American Culture*. Princeton, NJ: Princeton University Press.
- Maass, Arthur, Maynard M. Hufschmidt, Robert Dorfman, Harold A. Thomas, Jr., Stephen A. Marglin, and Gordon Maskew Fair. 1962. *Design of Water-Resource Systems: New Techniques for Relating Economic Objectives, Engineering Analysis, and Governmental Planning*. Cambridge, MA: Harvard University Press.
- Macchiavelli, Niccolò di Bernardo. [1532] 1979. "The Prince." In Bondanella, Peter, and Mark Musa (Eds. and Trans.) *The Portable Machiavelli*: 77–166. New York: Penguin Books.

- MacIntyre, Alasdair. 1984. *After Virtue*, 2nd Edn. Notre Dame, IN: University of Notre Dame Press.
- Marshall, Alfred. [1920] 1946. *The Principles of Economics*, 8th Edn. New York: Macmillan.
- McCarthy, Gina. 2019. "The Role Of Environmental Economics in US Environmental Policy." *Review of Environmental Economics and Policy*, 13(2): 299–307.
- McGartland, Al. 2021. "Quality Science for Quality Decisions: Protecting the Scientific Integrity of Benefit-Cost Analysis." *Review of Environmental Economics and Policy*, 15(2): 340–351.
- Moscati, Ivan. 2019. "Not a Behaviorist: Samuelson's Contributions to Utility Theory in the Harvard years, 1936–1940." In Anderson, Richard, William Barnett, and Robert A. Cord (Eds.) *Paul Samuelson: Master of Modern Economics*. London: Palgrave Macmillan.
- National Park Service (NPS), Land and Recreational Planning Division. 1949. *The Economics of Public Recreation: An Economic Study of the Monetary Evaluation of Recreation in the National Parks*. Washington, DC: National Park Service.
- Nussbaum, Martha C., and Amartya K. Sen. 1989. "Internal Criticism and Indian Rationalist Traditions." In Krausz Michael (Ed.) *Relativism, Interpretation, and Confrontation*: 229–325. Notre Dame, IN: Notre Dame University Press.
- Pigou, A.C. [1932] 1962. *The Economics of Welfare*, 4th Edn. London: MacMillan & Co.
- Plato. 1924. *Laches*, v. II of *Plato*. Translated by W. R. M. Lamb. Cambridge, MA: Harvard University Press.
- Porter Theodore, M. 1995. *Trust in Numbers: The Pursuit of Objectivity in Science and Public Life*. Princeton, NJ: Princeton University Press.
- Putnam, Hilary. 2002. *The Collapse of the Fact/Value Dichotomy and Other Essays*. Cambridge, MA: Harvard University Press.
- Putnam, Hilary, and Vivian Walsh (Eds.) 2012. *The End of Value-Free Economics*. London: Routledge.
- Reiss, Julian, and Jan Sprenger. 2020. "Scientific Objectivity." In Zalta, Edward N. (Eds.) *The Stanford Encyclopedia of Philosophy*. Available at <https://plato.stanford.edu/archives/win2020/entries/scientific-objectivity/>.
- Robbins, Lionel. 1935. *An Essay on the Nature and Significance of Economic Science*. 2nd Edn. London: Macmillan.
- Robbins, Lionel. 1938. "Interpersonal Comparisons of Utility: Comment." *Economic Journal*, 48: 635–641.
- Samuelson, Paul A. 1950. "Evaluation of Real National Income." *Oxford Economic Papers*, 2(1): 1–29.
- Scarantino, Andrea. 2009. "On the Role of Values in Economic Science: Robbins and His Critics." *Journal of the History of Economic Thought*, 31(4): 449–473.
- Scitovsky, Tibor. 1941. "A Note on Welfare Propositions in Economics." *Review of Economic Studies*, 9(1): 77–88.
- Sen, Amartya. 1967. "The Nature and Classes of Prescriptive Judgments." *The Philosophical Quarterly*, 17(66): 46–62.
- Sen, Amartya. 1987. *On Ethics and Economics*. Oxford: Basil Blackwell.
- Sen, Amartya. 1992. *Inequality Reexamined*. Cambridge, MA: Harvard University Press.
- Smith, Mark C. 1994. *Social Science in the Crucible: The American Debate over Objectivity and Purpose, 1918–1941*. Durham, NC: Duke University Press.
- Su, Huei-Chun, and David Colander. 2013. "A Failure to Communicate: The Fact-Value Divide and the Putnam-Dasgupta Debate." *Erasmus Journal for Philosophy and Economics*, 6(2): 1–23.
- Sugden, Robert. 2009. "Can Economics be Founded on 'Indisputable Facts Of Experience'? Lionel Robbins and the Pioneers of Neoclassical Economics." *Economica*, 76: 857–872.
- Sunstein, Cass R. 2023. "Inequality and the Value of a Statistical Life." *Journal of Benefit-Cost Analysis*, 14(1): 1–7.
- US Office of Management and Budget (OMB). 2023. *Circular A-4*. Draft for Public Review. Available at <https://www.whitehouse.gov/wp-content/uploads/2023/04/DraftCircularA-4.pdf>.
- Walsh, Vivian. 2003. "Sen after Putnam." *Review of Political Economy*, 15(3): 315–394.