

Cost-Benefit Analysis, Ethical Values, and a "Taste" for Fairness¹

By Patricia Marino

Abstract:

A challenge for cost-benefit analysis is that it ignores ethical values such as justice, fairness, and equity. One standard response is to regard CBA results as just one factor in a more complex decision-process where ethical and democratic factors are also considered. This paper considers an alternative response: extending CBA so that it takes into account not only self-interested input but also moral preferences such as a "taste" for fairness. Drawing on existing research and the example of resource allocation, the paper develops and analyzes objections to extended CBA. Evaluation of these objections, the paper shows, depends on how the original challenge is interpreted and how the problem of "ignoring" ethical values is understood. While some interpretations lead to an impasse between defenders and critics of extended CBA, the paper proposes a novel interpretation -- focused on political representationality -- and showcases the limits of CBA as a coherent response.

Introduction

In discussions of normative economics, policy, and law, cost-benefit analysis (CBA) and related methods are said to face the problem that efficiency-based criteria "ignore fundamental ethical norms" (Zamir and Medina 2019). As Acland says, the "textbook" definition of CBA is

¹ I presented versions of this paper at the 2022 NYU Philosophical Bioethics Workshop, the 2022 Legal Theory Workshop, the 2023 meeting of the Canadian Philosophical Association, the 2023 meeting of the Philosophy, Politics, and Economics Society meeting, the 2023 Rocky Mountain Ethics Conference, and the 2023 meeting of the International Network for Economic Methodology; thank you to all participants for their questions and discussion. I am also grateful to Mark Fabian, Rebecca Stone, Brian Talbot, and Hasko von Kriegstein for their helpful comments on earlier drafts.

that it deploys the "Kaldor Hicks" criterion that "a policy should be implemented if those who gain from the policy could compensate those who lose and still be made better off" (Acland 2022). "Gain" and "lose" here are standardly interpreted as referring to a person's individual welfare or self-interested preferences only (Adler and Posner 2006). This, together with its consequentialist nature, means that CBA typically leaves out considerations regarding distribution and fairness, constraints against harming others, and other ethical values.

One standard response is to adopt "weak welfarism" or "moderate deontology," in which conclusions about efficient outcomes are just one factor considered among others in a decision process (Adler and Posner 2006; Zamir and Medina 2019). However, an alternative approach is to extend efficiency-based reasoning so that "benefits" and "costs" include not only welfare-based preferences but also moral or value-based preferences such as a "taste for fairness" (Kaplow and Shavell 2009). In an overview, Zamir and Medina (2019) say that this approach is "more popular" than moderate deontology, at least among "lawyer-economists."

As Adler (2014) emphasizes, such an approach is non-standard, because insofar as the use of CBA is justified by appeal to the intuitive value of maximizing well-being, it ought to take into account only preferences linked to welfare in the right kind of way; insofar as the satisfaction of moral preferences does not make us individually better off, they are the "wrong kinds of preferences." Contra this view, however, Posner and Sunstein (2017) argue that satisfaction of moral preferences generally *does* make a person better off; for this reason, they have an individual utility aspect which ought to be accounted for in CBA. Going beyond moral preferences, Zerbo suggests that "all goods should be included in economic analysis for which there is a willingness to pay (WTP)" (2004; see also 2002). As Acland says, given the wide range

of outcomes people are willing to pay for, "This would be a dramatic change in the consensus approach to [CBA]" (2021, 3).

Proposals for "extended CBA" have not been the subject of extensive philosophical attention. Here I draw on existing critiques of several such proposals to analyze extended CBA as a response to the problem that CBA ignores ethical values. I argue first that the problem can be interpreted in multiple ways: as a substantive moral one, in which non-consequentialist values are just as weighty as considerations of well-being in decision-making so must not be ignored, as a political one, invoking an assumption that public decision-making should be appropriately representational, or as a narrowly methodological one, where we improve the counting of overall well-being through extension. I then consider three challenges to extended CBA: 1) that it misunderstands the relationship between moral preferences and moral judgments, 2) that it lacks a coherent foundation for aggregation, and 3) that including altruistic preferences can lead to inequitable results. I show that while none creates deep difficulties for methodological extensions of CBA, each forms a substantive challenge to extended CBA in its moral or political guise. However, I also argue that no proponents of extended CBA are concerned with the moral problem. Therefore, the political interpretation not only shows deep difficulties for extended CBA, it is also the only interpretation forming a site for genuine disagreement among participants in the debate. My contribution thus frames the problem of CBA and ethical values in a new and interesting way and also shows the limits of extended CBA as a way of addressing that problem.

CBA and ethical values: interpreting the problem

CBA is a family of related approaches based on "the numerical weighing of advantages against disadvantages" (Hansson 2007). Options are compared by considering their consequences, often by assigning them a monetary value. One common methodology uses contingent valuation, where people are asked what they are willing-to-pay (WTP) to for an outcome. CBA methods can be used as decision-rules -- processes in which the formal result becomes the decision -- or as only part of a decision-process, to be considered alongside other considerations in a more open-ended decision process e. g., in the weak welfarism mentioned above. CBA methods are typically deployed in the context of public decision-making including particular decision situations, such as when a government uses CBA to evaluate a proposed policy or regulation, and CBA reasoning is used in the law and economics context when we evaluate policies by considering their efficiency.

As Hansson explains, the family of methods in the CBA category can be seen as "applied consequentialism" -- that is, as manifestations of ethical views such as utilitarianism, in which the right action is the one that maximizes utility (2007, p 164). As with utilitarianism, objections often concern the possibility that the theory ignores ethical values such as justice, fairness, and equity. These objections have been characterized in several ways. For Zamir and Medina, the problem is that "the criteria of economic efficiency tend to ignore fundamental ethical norms" -- such as the inherent immorality of deliberately harming other people (2019, 24.1). Lowry and Peterson frame the issue by saying that CBA is thought to be "incompatible" with moral theories that are not utilitarian -- "particularly those that focus on deontological concepts such as rights (2011, 258)." Sagoff (1981) and others argue that while CBA can take into account consumer preferences, which are self-interested, it doesn't include citizen values, which may reflect

concern for society. Zerbe criticizes traditional CBA for ignoring "regard for others," especially as manifested in our concern for items such as fairness, waste, and income distribution (2002, 24).

One interpretation of the problem is a substantive moral one: non-consequentialist concerns such as fairness and respect for rights are, in fact, just as weighty as considerations of overall wellbeing and so must not be "ignored" in policy making. For example, Zamir and Medina (2019) argue that the consequentialist foundations of the economic approach to law may lead to the conclusion that killing one innocent person to save two others would be justified. But this violates an ethical principle against killing. Similarly, a tax policy supported by CBA reasoning might lead to forms of inequality that violate the requirements of distributive justice.

While this framing shows the core of the difficulty, it entails that the power of the objection rests on assumptions about either metaethics -- e. g., that there are moral facts -- or normative ethics -- e. g., that the correct moral theory is a non-consequentialist one such as deontology, in which killing one person to save others is wrong. In contexts of moral diversity, we also face the problem of how can we determine which "fundamental ethical norms" deserve consideration and why.

I argue that there is a closely related second interpretation of the problem that avoids these complexities. In what I call a "political" interpretation, the problem is that decision-making ignores the ethical values of the people whose interests CBA is meant to reflect, and thus fails to be properly representational. We may make this idea more precise by using the economics concepts of "social values" and "standing." Social values are "preferences or evaluative claims about an aggregate or community of persons" (Menzel 1999) -- where social values are preferences, I'll call them "social value preferences"). Social values thus reflect ethical values.

"Standing" refers to the specification of whose costs and benefits are being considered -- for example, this could be the citizens of a country or the stakeholders in an organization (Boardman et al. 2018). In the political interpretation, the idea would be that insofar as we accept that public decision-making should be appropriately representational or democratic, and insofar as social values reflect part of what people care about, those values should be reflected, not "ignored," in a decision-making process. For example, in evaluating a tax policy, social values might track ethical beliefs and attitudes about the value of equality and when and how inequality is wrong. These are ignored – and thus not represented -- in standard CBA, which takes into account only self-interested preferences. While the name "political" may seem to suggest governmental contexts, this interpretation is relevant anywhere decisions on behalf of a group are meant to represent the input and perspectives of that group. By contrast to the moral interpretation, in the political interpretation values such as equality and justice matter only derivatively -- because representation matters -- and not because they are true values in a metaethical sense.

As we will see in more detail in the following section, the moral and political interpretations of the problem contrast with one that is more narrowly "methodological" -- namely, that since CBA is about maximizing welfare, insofar as satisfying people's social preferences does increase their welfare, those preferences should not be "ignored." If people are willing to pay to have their social value preferences satisfied, we may interpret that willingness as reflecting an increase in their well-being corresponding to those preferences being satisfied. It would follow that considering only self-interested preferences leads to undercounting the benefits of enacting a particular proposal. It would then be a methodological improvement to include social value preferences, including when those preferences reflect ethical values. For example, in evaluating a tax policy, if people have social value preferences about inequality in

their society that they are willing to pay to have satisfied, the satisfaction of those preferences is a well-being gain for them, so a CBA should take such preferences into account and not "ignore" them.

To explore these interpretations of the problem of CBA and ethical values, consider the context of health care resource allocation. There, one standard and often criticized methodology uses cost-effectiveness analysis (CEA) – a member of the family of CBA methods in which benefits are counted in a non-monetary metric. In this context, benefits are often expressed in terms of quality-adjusted life years (QALYs), which measure the effectiveness of an intervention through health and longevity, with the burdensomeness of a health state evaluated on a scale from 0 (worst) to 1 (best) so that a year of full health is one QALY. In evaluating a treatment, an organization might consider how many QALYs it is likely to produce per unit cost. With such a calculation, an organization may fund treatments up to a fixed budget or may use a "threshold" model, funding treatments to a given cost-per-QALY (Brazier et al. 2016). In a threshold model, WTP surveys are also used to estimate what cost-per-QALY people are prepared to pay (Bertram et al. 2016).

These methods are known to raise well-known difficulties (Bickenbach 2021). CEA can lead to discrimination against people with disabilities, as they are often judged to have a quality of life lower than that of able-bodied people, so that interventions extending their lives are seen as less effective (Brock 2004; Bogner 2010, 2011). Since "a QALY is a QALY," the CEA process is insensitive to questions of distribution and equity, with no priority for younger people or the worse off and no amelioration of existing health inequities, even when those inequities are themselves rooted in oppression or poverty.

These tensions highlight the difference between citizen preferences about individual health states (e. g., QALY evaluations) and those related to how health states should be handled in a system (e. g., what to fund and prioritize; see Helgesson 2022). The former are self-interested preferences and the latter are social value preferences reflecting ethical values such as those of equality and justice. The moral interpretation of the problem would be that CEA methods of resource allocation violate actual norms of justice and equity. The political interpretation of the problem would be that CEA methods fail to reflect citizen preferences about justice and equity and how those values would inform resource allocation decisions. The methodological interpretation of the problem would be that in considering only self-interested benefits, and not the satisfaction of social value preferences, the method fails to count benefits accurately. I say more about these interpretations in the following sections.

Some extended CBA proposals and their motivations

One standard response to problems of CBA and ethical values is to consider CBA as only one component in policy-making: instead of using it as a decision rule, we may adopt "weak welfarism," where the results of CBA are just one factor to be balanced among others in a decision process. This response aligns with the analysis given in Boardman et al.'s (2018) standard CBA textbook: in response to concerns that CBA is "undemocratic," they say that CBA rarely serves as a "decisive decision rule." In his extensive work on CBA, Adler urges that the benefits counted in CBA should reflect only individual well-being and that CBA should thus not be used as a decision-rule: when it comes to decision-making, we should bring in other values (Adler and Posner 2006; Adler 2014). Acland (2021) says that while US government policy is

vague, we cannot use CBA as a decision rule because "other values simply must be taken into account" (p. 3).

However, especially in the law and economics context, Zamir and Medina say that the more "popular" strategy is to "treat moral judgments as preferences whose fulfillment should be maximized along with any other" (2019, section 24.4.1) in what I am calling "extended CBA." Citing in part the work of Kaplow and Shavell, Posner and Sunstein, and Zerbe, Zamir and Medina describe extended CBA as attempting to "incorporate people's moral attitudes -- including the preference for fairness and prohibitions against harming others -- into CBA."

I start by giving brief overviews of proposals from these authors and their motivations. In their book *Fairness versus Welfare*, legal theorists Kaplow and Shavell (2002) argue that only welfare considerations deserve direct consideration in policy-making; in particular, fairness should be accorded no "independent weight." They say that not only can fairness perversely affect welfare, fairness judgments are subjective, personal, and best regarded as the misleading result of evolution. However, welfare considerations are based on preference satisfaction, and moral preferences -- such as a "taste" for fairness -- should be treated the same way as any other preferences, and thus taken into account alongside self-interested preferences.

Their reasoning in favor of including moral preferences is based on methodological neutrality. Especially given a diversity of moral views, they argue, policy should be based on a framework that is appropriately objective. "Neutrality," they argue, requires that we base policy on what actually matters to people. Moral preferences should therefore be treated the same way as any other preferences: to leave them aside would be "arbitrary." "Just as an individual might derive pleasure from art, nature, or fine wine," they explain, "so might an individual feel better

with the knowledge, for example, that vicious criminals receive their just deserts" (2002, 333).²

Posner and Sunstein (2017) base their argument for taking moral preferences into account on the idea that their satisfaction generally does make a person better off, and conversely; moral preferences have an individual well-being aspect that ought to be accounted for in CBA. For example, if a person is willing-to-pay to protect an endangered species, then they are made better off when that species is protected; what they are willing to pay reflects the degree to which their well-being is increased. Focusing on US regulatory requirements, they argue that to focus on self-interested preferences only risks undercounting the benefits a policy can bring, which is especially worrisome when CBA is used to block regulations. If we have preferences for income equality, for reducing prison rape, or for protecting the environment, then excluding moral preferences fails to count benefits correctly.

Like others, Zerbe defends CBA partly on grounds of consistency -- that we should treat all WTP preferences the same way -- and that the aim of CBA is to reflect "what the public wants" (2002, p. 17). Since much of what the public wants reflects their concerns for others, "including the welfare of people to whom they are not related," (2002, p.26), CBA ought to treat moral preferences in the same way as other preferences. In contrast to others, Zerbe also refers to considerations of pragmatics and the norms of policy-making. When CBA ignores the care and regard for others that people generally have, he argues, it will fail to reflect people's full range of concerns, which will make it less appealing to people and less useful. Extended CBA better instantiates the ideal of a "social contract" between people and policy-makers.

² Kaplow and Shavell use the more restrictive Pareto criterion rather than the KH standard. See Coleman 2003 for discussion.

To see the commonalities and differences among these proposals, notice first that in all three, preferences concerning moral values are included in the CBA: when those preference are satisfied, this is a benefit counted in the same way as the satisfaction of an individual self-interested preference. Posner and Sunstein give the example of assessing a regulation to require companies to disclose their use of "conflict minerals": minerals mined where armed groups profit while contributing to violence and exploitation. If people feel a moral obligation not to use products that contain conflict minerals, they will benefit from receiving disclosure information. From a practical perspective, the three proposals for extending CBA recommend a similar way of proceeding in such a case: consider what people are likely to be willing to pay to receive such information and include that estimate among the benefits in the analysis.

With respect to why such preferences should be included, all authors refer to the idea that policies should properly reflect what people care about, as expressed through their preferences. As "social values" reflect what people care about, the logic of all these versions of extended CBA is aptly expressed using this concept instead of the slightly narrower one of "moral" preferences. Note that all three proposals consider preferences only. Depending on what views of metaethics or moral psychology are in the background, it is possible that a person could have ethical commitments or values that are not reflected in any WTP preferences or could have a strong moral commitment where there is only a weak WTP component. Since "social values" are defined as preferences or "evaluative claims," extended CBA thus extends not to moral judgments, and thus not to all social values, but rather just to preferences -- what we might call "social value preferences." While not all proponents are explicit on this matter, I assume here in line with Kaplow and Shavell that extended CBA is intended to be used as a decision rule. Because of these commonalities, I consider "extended CBA" to be any CBA that aims to include

the satisfaction of (some or all) of people's social value preferences among the relevant costs and benefits, where these may reflect moral or ethical preferences.

While all authors refer to the idea that policies should properly reflect what people care about, as expressed through their preferences, they give different interpretations of this idea: Kaplow and Shavell focus on neutrality and non-arbitrariness, Posner and Sunstein on accurate counting of self-interested well-being, and Zerbe on the idea that CBA reflects an implied social contract. We will see below that these various motivations can lead to practical differences with respect to what gets counted and why.

Recall that in the moral interpretation, the problem of CBA and ethical values is that concerns such as fairness and respect for rights are just as weighty as considerations of overall wellbeing and so must be considered in policy making. Notably, none of the authors is responding to, or motivated by, the moral version of the problem. Kaplow and Shavell argue explicitly against including non-consequentialist considerations in policy; Posner and Sunstein say that including WTP evaluations reflects only a well-being component of a moral preference - only the "positive psychological effect" of satisfying a moral preference; Zerbe refers not to what matters morally but rather to what people believe and care about.

Recall next that the political version of the problem is that CBA is wrong to "ignore" ethical values because in doing so, we ignore part of what people care about, violating norms of representational decision-making. Zerbe, in his claim that CBA is based on a metaphorical social contract and should thus incorporate people's ethical values, including justice and concern for others, can be interpreted in light of the political interpretation of the problem of CBA and ethical values. Ignoring ethical values expressed in people's moral preferences in CBA is wrong because doing so renders CBA less able to represent the concerns of the relevant people.

The motivations of Kaplow and Shavell and Posner and Sunstein are best understood not as motivated by moral or political concerns but rather by more narrowly methodological considerations. They propose extended CBA simply as a better, more accurate, and more principled way of doing CBA. When ethical values are expressed in people's preferences, it is wrong to ignore them because it is arbitrary to fail to include them (Kaplow and Shavell) or because including them leads to a better methodology for counting welfare (Posner and Sunstein). Thus, insofar as they may be interpreted as responding to a problem of CBA and ethical values, that problem would be the methodological one described above: CBA is wrong to "ignore" ethical values because including preferences based on them would improve our accounting of individual well-being.

To sum up: There are three interpretations of the problem of CBA and ethical values: a moral one, a political one, and a narrowly methodological one. Kaplow and Shavell and Posner and Sunstein are best interpreted in light of the methodological interpretation; Zerbe in light of the political one. None is concerned with or motivated by the moral version of the problem. Next, I consider three challenges to extended CBA, and I argue that whether and how these lead to problems for its use depends on the motivation underlying the extension.

Challenges for extended CBA 1: judgments versus preferences

One prominent criticism of extended CBA emphasizes the distinction between moral judgments and preferences. In a discussion of Kaplow and Shavell (2002), Kornhauser (2003) argues that moral considerations must be reflected through beliefs and judgments, not preferences, where judgments and preferences are crucially different. Moral beliefs are not like personal preferences concerning chocolate or vanilla because 1) unlike moral judgments,

preferences are "personal claims" which do not carry implications for what others should do and

2) people are sovereign over their preferences in a way they are not over their judgments (2003, 317-318). Judgments, Kornhauser says, should be reasoned about, not aggregated (for discussion see Zamir and Medina 2019, 24.4.2). Aggregating judgements about justice and fairness, he says, misunderstands the nature of moral judgment. To support this, Kornhauser alludes to Dworkin's (1986) example in which people disagree about the moral permissibility of abortion, but everyone objects to a scheme in which a proportional group of people fall under a liberal rule and a proportional group under a restrictive one, with people randomly slotted into group -- a scheme that would reflect satisfying "preferences" about whether abortion should be permitted. Referring to various critics, Zamir and Medina sum up by saying extended CBA wrongly treats "moral judgments as preferences whose fulfillment should be maximized along with any other," because "whether a judgment is right or not depends on its justifications, not on the number of its supporters or the intensity of its adherents" (Zamir and Medina 2019).

As it is formulated, this criticism misses its target. The criticism is that extended CBA treats judgments as preferences. But as we've seen above, the strategies offered by proponents of extended CBA do not treat moral judgments as preferences: all include moral preferences and leave judgments aside, seemingly respecting a distinction between them.

Consideration of context and motivation helps to clarify matters. I argue first that the criticism is irrelevant to extended CBA as motivated methodologically, because there, including preferences and leaving judgments aside is apt. If the point of extending CBA is to be consistent and neutral with respect to other preferences, or to count the well-being component of moral preferences accurately, then extended CBA should include preferences and leave judgments

aside. Understood this way, extended CBA is not concerned with judgments at all, so does not elide distinctions between moral judgments and preferences.

If the problem of CBA and ethical values is the moral one, then it is clearly relevant if CBA aggregates preferences rather than considering reasons and justification for moral claims, but in this case, the criticism would have to be reformulated: it is not that extended CBA treats moral beliefs as preferences, but rather that extending CBA is inapt for the purpose of respecting actual non-consequentialist considerations, because it is implausible that social preferences are information about the actual moral facts. That is: moral judgments should be based on reasoning and evidence, not majority rule, so canvassing preferences about moral matters would be mistaken. This line of thought strongly suggests that no version of extended CBA would adequately address the problem in its moral interpretation, which in turn shows the limits of how extended CBA should be used and what problems it can solve. Of course, since no proponents of extended CBA are engaged with the moral version of the problem, this is not a criticism of any of their specific proposals.

The political version of the problem shows a context where the criticism of judgments versus preferences not only engages a genuine disagreement but also shows a strong criticism of actual extended CBA. In this case, again, the criticism must be reformulated: it cannot be that extended CBA fails to recognize the distinction between judgments and preferences, as all versions include only preferences and not judgments. To frame a proper formulation, consider that underlying the criticism seems to be that to take into account moral values, public decision-making ought to reflect judgments -- so that reflecting only preferences is wrong. For example, if a person judges that a certain pattern of distribution is unjust, representationality should take that judgment into account. The political context allows us to make the logic explicit: 1) public

decision-making ought to be representational or democratic; 2) properly representational decision-making should reflect moral judgments, and not only moral preferences, 3) extended CBA takes into account preferences only so 4) it is an objection to extended CBA that it ignores moral judgments.

Though Kornhauser is not obviously focused on the political interpretation of the problem of CBA and ethical values, the preceding diagnosis is a good fit with his use of the abortion example above: it is because many people have moral judgments about abortion that policies reflecting only their preferences will be seen as failures. The political interpretation of the significance of judgments versus preferences supports Kornhauser's claim that the preferences in extended CBA are not an "appropriate maximand" for a policy maker (p. 209), but diverges in explaining why this matters. It matters not because extended CBA is aggregating moral judgments (because it is not), but rather because it wrongly leaves them aside.

In support of this form of the criticism of preferences versus judgments, note that in the resource allocation context, one recent attempt to properly reflect social values in a decision process asks citizens for input on how collective resources should be allocated – and crucially, asks for their judgments, not preferences (Richardson et al., 2014; Ubel et al. 2000). That is, instead of soliciting preferences, this alternative to CEA asks people to consider their judgments – not preferences -- about how a pooled resource, such as government funds, should be distributed. This example supports my claim that the preferences versus judgements challenge for (political) extended CBA is that it wrongly includes only the former where the latter are methodologically relevant.

Overall, then, the problem of judgments versus preferences does not apply to extended CBA as motivated methodologically but only as motivated morally or politically. In the political

context, it arises because extending CBA does not achieve a relevant aim: representation would require reflecting ethical and evaluative judgments, but extended CBA does not.

Challenges for extended CBA 2: the problem of apples and oranges

Another challenge for extended CBA focuses on difficulties of aggregation. We've seen a critique of traditional CBA that because it takes into account "consumer preferences" but not "citizen" or "societal" values it is an unsound basis for public decision-making (Sagoff 1981; Anderson 1995). Extended CBA could be seen as addressing this difficulty.

However, aggregating various kinds of preference can lead to incoherence. One issue is that "social value preferences" can be from two different points of view: a self-interested, personal one or an ethical, societal one, and these perspectives can lead to different preference rankings. For example, a person may prefer less inequality from an ethical point of view and greater inequality from a self-interested point of view. In resource allocation, a person might prefer to fund treatments for rare diseases from an ethical point of view, but not from a self-interested point of view. The environmental economist Nyborg points out that in answers to WTP questions, it may be unclear which perspective people are deploying -- especially as many surveys are ambiguous. But preferences from the self-interested point of view and preferences from the citizenship view are so different, she says, that the problem of combining them is one of adding "apples and oranges" (2000, 319).

The sense in which there are potentially incompatible perspectives may seem unclear. Defenders of extended CBA may point out that it is a standard idealization to assume that people have one set of consistent preferences: if a person wants contradictory things, that's a problem with the agent's rationality, not with cost-benefit analysis (see Ratterman 2007). But there are at

least two distinct ways of drawing distinctions among social preferences that do not hinge on agent irrationality. One is that preferences from the self-interested perspective address the question of benefits for the individual while preferences from a societal perspective address the question of "community benefits to society" (Howley et al. 2010; see Nyborg, 2000). Economists point out that confusion can be exacerbated by survey design that shifts away from a "pseudo-market" setting toward a "political choice setting": the latter may elicit decisions based on social welfare rather than self-interested preferences (Blamey 1995, 263, cited in Ami et al. 2014). Surveys may try to lessen confusion by framing WTP questions with "Bearing in mind the importance or unimportance of ... for you personally" or "Bearing in mind the importance or unimportance of ... for society as a whole" (Howley et al. 2010). The point here however is that the two kinds of preferences are fundamentally different, as one measures benefits to an individual and the other measures benefits to a group.

A second distinction is one Nyborg (2000) frames between "sole responsibility" WTP evaluations, which address the question of how much an individual is willing to pay themselves, and "shared responsibility" WTP evaluations, which address the question "What is the maximum amount I find it socially right for everybody to pay, in order to ensure this project?" Arguing that survey questions can be interpreted to mean either, she notes empirical research suggesting that those using a "shared responsibility" interpretation give lower valuations, partly because "if everyone did their part, each household would not have to give all that much" (2000, 313). Nyborg argues that with ambiguous surveys questions, we may not know which respondents are using. But, again, what is relevant here is that even if the questions are clear, "sole responsibility" and "shared responsibility" evaluations concern different things.

Neither distinction requires attributing any agent irrationality. But both show how social value preferences can concern different objects of evaluation. The resource allocation context helps to illustrate the differences. QALYs attempt to measure self-interested preferences. Social value preferences are preferences about how health states are handled or treatments are prioritized, and these could be from a self-interested or societal perspective. Imagine we are deciding between a treatment that provides four people with three QALYs each or an equally costly one that provides three of those four people with five QALYs each and the fourth person with nothing (see Richardson and Schlander (2019)). Counting QALYs, the first treatment is most cost-effective. But in social values from a societal perspective, we might prefer the second, as it is more equitable. From a self-interested perspective, our social value preferences could go either way: a risk-averse person may not want to risk being that fourth person with nothing, while another might prefer to play the odds, depending also on a person's estimate of their personal likelihood of being affected by the health state in question. When WTP input is used to set a cost-per-QALY threshold, "individual values" are what people are willing to pay to improve their own health, whereas the "social value" of a QALY is defined as what people are willing to pay to "contribute to a health gain achieved in society" (Bobinac et al. 2013). Adding further complexity, a "sole responsibility" judgment might reflect what a person is willing to pay per QALY for themselves while a "shared responsibility" judgment might reflect what a person would be willing to pay per QALY in a system where everyone pays the same amount.

The potential problem for extended CBA concerns the possibility of coherent integration of these different kinds of preferences -- adding apples and oranges. Notice, however, that this question of integration and adding apples and oranges does not arise for methodological extensions of CBA such as those of Kaplow and Shavell and Posner and Sunstein. If extended

CBA is intended to reflect the well-being component associated with the satisfaction of a moral preference, then what is being counted is the same in all cases: the individual welfare associated with the preference being satisfied. This is the preference relevant only to "Bearing in mind the importance or unimportance of ... for you personally" questions. And in this context, presumably we are interested only in "sole responsibility" evaluations and not "shared responsibility" ones. There are challenges in designing survey questions, but conceptually there is one quantity being aggregated. This is adding apples and apples.

If the motivation for extended CBA were to address the moral problem, then plausibly we would have an apples and oranges problem because even setting aside the problem of judgments versus preferences, preferences from a self-interested perspective and from a moral perspective address two different kinds of consideration. As no proponents of extended CBA are in this category, it is not a criticism of any specific proposal.

I argue that the political setting, using the concepts of social values, enables precision about the criticism and shows how it applies. "What people care about" includes both individual and societal perspectives and is reflected in both sole responsibility and shared responsibility judgments. But how can we integrate these fundamentally different kinds of considerations in a principled way? Given that self-interested personal preferences and self-interested social value preferences both concern evaluations of good for one's self, and social value preferences from a societal perspective concern evaluations of good for society, any way of integrating them would face questions of how they compare and what trade-offs among them are appropriate. These questions hinge partly on how public decision-making should balance individual welfare and collective good. Since this is one of the most contentious questions in political philosophy, there is no obvious or uncontroversial answer. For sole responsibility and shared responsibility

judgments as well, if we know that people are willing to pay X dollars for an outcome as an individual and Y dollars if everyone else is paying Y dollars, how should this information be usefully combined in a quantitative way? There is nothing particular we are aggregating if we add up the Xs and Ys together.

In traditional CBA, there is assumed to be a quantity we are quantifying and aggregating: individual well-being. But there is nothing analogous answering to the foundation for integrating individual well-being and societal well-being, or for trying to combine sole responsibility and shared responsibility social value evaluations. For extended CBA motivated by the political problem of representing what people care about, the challenge is that even taking into account only preferences (and not judgments), we have no way to integrate quantitative information reflecting measurements of different quantities in a principled way.

Overall, then, the problem of apples and oranges does not apply to extended CBA as motivated methodologically but only as motivated morally or politically. In the political case, it arises because reflecting the full range of what people care about would require taking into account fundamentally different kinds of input -- such as preferences from an individual perspective and those from a societal perspective -- where there is no consensus on how these forms of input could be integrated.

Challenges for extended CBA 3: altruism, equity, and fairness

Extended CBA also faces potential challenges arising from the inclusion of altruistic preferences. From an economic theory perspective, one characterization of altruism involves one person's utility function depending on that of another. Altruistic preferences can be understood as a subset of "other-regarding" preferences which can also include negative preferences based on

envy, hatred, and so on. Recall that Zerbe's extension of CBA means including preferences based on "care for others." It is well-known that preferences can reflect altruism: a WTP determination may reflect not only the benefits that a person expects to experience themselves but also the benefits that they expect others to experience. Insofar as they care about others, they will factor this into their WTP determination (Milgrom 1993, Bergstrom 2006).

To see why altruism may be a concern for aggregation, consider problems familiar from social contract theory regarding the way that altruistic preferences lead to bargaining outcomes that may seem inapt or unfair. Contractarian Gauthier says, "[Supposing that] I, considering us equally fond of cake, prefer that each of us get half, not only to your having a larger share but also to my having it, and [that] you prefer more cake for yourself to less, whatever I get, then it seems implausible to suppose that a rational and fair division gives you three-quarters of the cake and me one-quarter" (Gauthier 1986; see Morris 1991). The selfish person has had their preference double-counted.

Problems to do with altruistic preferences and double-counting are well-known in debates even over traditional CBA. One worry concerns double-counting leading to inaccurate assessment of benefits (Milgrom 1993; see also Dworkin 1978 on "extended preferences"). More directly relevant in the present context is the possibility that including altruistic preferences can lead to inequities. Milgrom (1993) and others point out that with aggregation, targets of altruism will end up better off than who are less likely to be the object of care or positive attention. "Incorporating concerns for others in benefit-cost calculations," Milgrom says, "would also tend to justify as 'efficient' projects and policies that favor more popular groups relative to less popular ones" (page 421). Policies benefitting those in groups less cared for -- such as those in marginalized communities -- will be less likely to be adopted than policies benefitting the well-

loved or appreciated. In resource allocation, some concerns over altruistic preferences manifest in debates over whether to include "spill-over effects" -- impacts on family members and carers - in the evaluation of a treatment's effects. Wittenberg (2019) et al. argue that while these effects do seem relevant as benefits and burdens when we are evaluating the effect of a given treatment, including them can lead to inequitable results, as interventions more likely to aid isolated patients, such as the unhoused, could be "undervalued" compared to those aiding people who are more loved and connected, such as children (p. 495).

While inequitable results for the less popular can arise in traditional CBA, it is a more challenging problem for extended CBA. The main reason is that extending CBA to include social value preferences suggests that any altruistic preferences must be included and aggregated with non-altruistic, self-interested ones. Debate over altruism in traditional CBA acknowledges the complexities in deciding to include altruistic preferences of some forms and not others. But once we leave aside the restrictions of traditional CBA, there is no principled way to exclude altruistic preferences. The idea of extending CBA to be neutral with respect to preferences and to reflect what people actually care about would support including altruistic preferences. And insofar as social value preferences concern our community or society, altruism is likely to be reflected in those preferences.

Analogously with the cake example above, CBA can lead to outcomes negatively affecting altruists. If the satisfaction of an altruistic benefit is counted as a benefit alongside self-interested preferences, and if only some people are altruistic, those toward whom they direct their altruism will have their self-interested preferences double-counted. Altruists will end up what might be called "materially" worse off: while their caring preferences are satisfied, they

have less of what they self-interestedly prefer, as compared to their self-interest gains had they been less altruistic or had others been more so.

Again, while double-counting can be a challenge for traditional CBA, the challenge is exacerbated for extended CBA, because extension means eliciting input not only on how an outcome or policy affects us as individuals but also on the outcome's effects on society. If altruism in a population is correlated with preferring certain policies or outcomes, those outcomes will be seen as less beneficial and thus less efficient compared to those favored by the self-interested. For example, imagine that in a given society that women are more altruistic overall than people of other genders, and suppose breast cancer is proportionately more likely to affect women. Insofar as extended CBA takes into account social value preferences about how health states are handled, if altruistic women express preferences that treatments for a range of health conditions are equally important and others express self-interested preference prioritizing treatments for health conditions they are more likely to have themselves, then breast cancer treatments will be less well-funded than they would be if only self-interested preferences were counted. Or imagine that younger people are more altruistic and older people are more self-interested. Then outcomes preferred by older people will be more likely to be selected for implementation. As in the cake example, where altruistic and self-interested preferences are aggregated, altruists end up materially worse off than they would if CBA had used self-interested preferences only -- that is, with respect to what they prefer for themselves, they end up worse off than they would be if they had been less altruistic or had others been more so.

Are these equity implications unfair? In the bargaining context, Gauthier takes the unfairness of cake distribution as obvious, and Gaus (2021) argues that if outcomes are worse for the altruistic because of double-counting of self-interested preferences, "someone who has

already taken into account the welfare of others is required to concede even more"; this is "exploitative" (see also Dimock 1999). On the other hand, it may be said that the results are appropriate even when they seem inequitable: if altruists are better off when their altruistic preferences are satisfied, then the effects of double-counting are not a problem. If altruists gain less of what they would have preferred for themselves, that is in accordance with their preferences, so perhaps this is "fair." And if the point of the CBA is to reflect what people care about, then perhaps popularity is a reasonable guide to priority.

I do not attempt to adjudicate this complex matter here; instead, let us return to the main question of how these matters bear on extending CBA. When extension is motivated methodologically, one possibility is to reject concerns over unfairness or exploitation as simply irrelevant. We've seen that Kaplow and Shavell's extension of CBA goes hand-in-hand with a rejection of fairness judgments as subjective, misleading, and unsuitable for use in policy. Adopting this perspective leads to a kind of consequentialism for policy, in which ethical values such as fairness or justice have no direct role to play in decision-making, so any fairness or exploitation considerations to do with double-counting are not relevant. This narrow methodological response renders extended CBA coherent: any intuitive judgments about the unfair effects of double-counting and unpopularity may be rejected as unreliable and beside the point.

As with the previous two challenges, from the point of view of the moral interpretation, we may say straightforwardly that double-counting is wrong insofar as it leads to unfairness and exploitation. This gives us further reason to believe extended CBA could not be a solution to the moral version of the problem, but again, no proponents are concerned with that version.

Also as above, the political perspective allows for a precise articulation of the difficulty and leads to genuine disagreement. With respect to whether the equity implications are unfair, exploitative, or otherwise wrong, this is itself a matter about which people may have social value preferences. As Fleurbaey et al. (2019) note in a discussion of other-regarding preferences and fairness, there is "a risk of inconsistency when the basic principles of social evaluation are not respected by the other-regarding preferences of the individuals." For example, other-regarding preferences may display a bias against a minority -- or, in terms of altruism, a bias in favor of a dominant social group. But this may violate norms of impartiality in decision-making processes. Should preferences for impartiality be factored in alongside the biased preferences that lead to the violating implications? Or handled in some other way? There is no clear answer to what it means for CBA to appropriately reflect what people care about when those cares include altruism, potential inequities, and social value preferences about unfairness.

Overall, then, problems related to altruism, equity, and fairness do not apply to extended CBA as motivated methodologically but only as motivated morally or politically. In the political context, they arise because including other-regarding preferences can lead to forms of unfairness or exploitation that conflict with principles of social evaluation such as impartiality or violate existing social value preferences regarding unfairness and exploitation; we have no methodology for how to handle such conflicts.

Conclusion

The problem that CBA ignores ethical values can be interpreted in multiple ways, leading to differential assessments when extended CBA is seen as being motivated by that problem. I have distinguished a moral interpretation, in which it is wrong to ignore actual non-

consequentialist considerations, a political one, in which representational public decision-making should reflect what people care about, including ethical values, and a narrowly methodological one, in which neutral or accurate accounting would include the well-being component of preferences based on ethical values. I've developed three challenges for extended CBA, and in each case, I've argued that while these may not apply in the methodological context, they lead to substantive criticisms when extended CBA is understood in light of the moral or political problem.

While no proponents of extended CBA are motivated by the moral version of the problem, reflections in this paper do imply no form of extension would succeed as a response: if the goal is to include actual non-consequentialist concerns such as fairness and respect for rights in policy-making, then challenges above show that extending CBA could not succeed. We would need alternatives such as weak welfarism or moderate deontology. Since at least some proponents of extended CBA reject non-consequentialist considerations on metaethical and normative grounds, this framing leads to an impasse: one side urges the role of ethical values and deontological considerations in public decision-making and the other rejects it.

However, our arguments show that the problem of CBA and ethical values is stronger than this impasse suggests. The political interpretation invokes an assumption that public decision-making should be appropriately representational: decisions should reflect what people care about, which includes ethical values and concern for others. I have argued that in this context, extended CBA fails as well; here, too, we would need alternatives such as weak welfarism or moderate deontology. Of course, the question of how public decision-making should function representationally is complex and multifaceted: for example, in a diverse society, whose values should prevail and why? Discussion of the three challenges above shows reason to

believe that extending CBA would be an unsuccessful way to address such questions. A topic for future research is whether there are other ways to extend CBA that would function better from a representationality perspective.

One way to deny the importance of representational decision-making is to insist, as some methodological defenders of extended CBA do, that only consequentialist values such as well-being matter in any case, so that regardless of the values of the people with standing, considerations such as fairness must never play a direct role in policy-making. Here, only the well-being aspect or component of moral preferences is taken into account: it is only when satisfaction of a moral preferences is a well-being increase that it factors into consideration. Decision-making is then fully consequentialist. I leave for another occasion evaluation of this way forward, which hinges partly on arguments in metaethics and normative ethics (see Coleman 2003).

References

- Adler, Matthew D., and Eric Posner. 2006. *New Foundations of Cost-Benefit Analysis*. Harvard University Press.
- Adler, Matthew D. 2014. "Extended Preferences and Interpersonal Comparisons: A New Account." *Economics & Philosophy* 30, no. 2: 123-162.
- Ami, Dominique, Frédéric Aprahamian, Olivier Chanel, Robert-Vincent Joulé, and Stéphane Luchini. 2014. "Willingness to Pay of Committed Citizens: A Field Experiment." *Ecological Economics* 105 (2014): 31-39.
- Anderson, Elizabeth. *Value in Ethics and Economics*. Harvard University Press, 1995.
- Bergstrom, Theodore C. 2006. "Benefit-Cost in a Benevolent Society." *The American Economic Review* 96, no. 1: 339-351.
- Bertram, Melanie Y., Jeremy A. Lauer, Kees De Joncheere, Tessa Edejer, Raymond Hutubessy, Marie-Paule Kieny, and Suzanne R. Hill. 2016. "Cost-Effectiveness Thresholds: Pros and Cons." *Bulletin of the World Health Organization* 94, no. 12: 925.
- Bickenbach, Jerome. 2021. "Disability and Health Care Rationing", *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), Edward N. Zalta (ed.), URL = <<https://plato.stanford.edu/archives/spr2021/entries/disability-care-rationing/>>.
- Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David L. Weimer. 2018. *Cost-Benefit Analysis: Concepts and Practice*. Cambridge: Cambridge University Press.
- Bobinac, Ana, N. Job A. van Exel, Frans FH Rutten, and Werner BF Brouwer. 2013. "Valuing Qaly Gains By Applying A Societal Perspective." *Health Economics* 22, no. 10: 1272-1281.
- Bognar, Greg. 2010. "Does Cost Effectiveness Analysis Unfairly Discriminate against People with Disabilities?" *Journal of Applied Philosophy*, 27(4): 394-408.
- Bognar, Greg. 2011. "Impartiality and Disability Discrimination," *Kennedy Institute of Ethics Journal*, 21(1): 1-23.
- Blamey, R., Common, M., Quiggin, J., 1995. Respondents to Contingent Valuation Surveys: Consumers or Citizens. *Austrian Journal of Agricultural Economics* 39, 263–288.
- Brazier, John, Julie Ratcliffe, Joshua Salomon, and Aki Tsuchiya. 2016. *Measuring and Valuing Health Benefits for Economic Evaluation*. Oxford University Press.
- Brock, Dan. 1995. "Justice and the Ada: Does Prioritizing and Rationing Health Care Discriminate against the Disabled?" *Social Philosophy and Policy*, 12(2): 159-185.

Brock, Dan. 2004. "Ethical Issues in the Use of Cost-Effectiveness Analysis for the Prioritization of Health Care Resources," in G. Khushf (ed.), *Handbook of Bioethics*, Dordrecht: Springer, 353-380.

Coleman, Jules L. 2003. "The Grounds of Welfare." *Yale Law Journal* 112.

Dimock, Susan. 1999. "Defending Non-tuism." *Canadian Journal of Philosophy* 29, no. 2: 251-273.

Dworkin, Ronald, 1978. *Taking Rights Seriously*. Cambridge, Mass.: Harvard University Press.

Dworkin, Ronald. 1986. *Law's Empire*. Cambridge, Mass.: Harvard University Press.

Fleurbaey, Marc. 2019. "Economic Theories of Justice." *Annual Review of Economics* 11: 665-684.

Gaus, Gerald. 2021. "Is Mutual Advantage a General Theory of Justice? More Domain Worries." *Philosophical Studies* 178, no. 5: 1731-1739.

Gauthier, David. 1986. *Morals By Agreement*. Oxford University Press.

Hansson, Sven Ove. 2007. "Philosophical Problems in Cost-Benefit Analysis." *Economics & Philosophy* 23, no. 2: 163-183.

Helgesson, Gert, Olivia Ernstsson, Mimmi Åström, and Kristina Burström. 2020. "Whom Should We Ask? A Systematic Literature Review of the Arguments Regarding the Most Accurate Source of Information for Valuation of Health States." *Quality of Life Research* 29, no. 6: 1465-1482.

Kaplow, Louis, and Steven Shavell. 2002. *Fairness Versus Welfare*. Harvard University Press.

Kornhauser, Lewis A. 2003. "Preference, Well-Being, and Morality in Social Decisions." *The Journal of Legal Studies* 32, no. 1: 303-329.

Lowry, Rosemary, and Martin Peterson. 2012. "Cost-Benefit analysis and Non-Utilitarian Ethics." *Politics, Philosophy & Economics* 11, no. 3: 258-279.

Milgrom, P. 1993. Is Sympathy an Economic Value? Philosophy, Economics, and the Contingent Valuation Method, In: Hausman, J.A. (Ed.), *Contingent Valuation. A Critical Assessment* (North-Holland: Amsterdam/New York)

Morris, Christopher. 1991. "Moral Standing and Rational Choice Contractarianism." In Peter Vallentyne (ed.). *Contractarianism and Rational Choice: Essays on David Gauthier's Morals by Agreement*. Cambridge: Cambridge University Press, pp. 76–95.

Preprint. Published in the *Journal of Economic Methodology*, 2024.

Nyborg, Karine. "Homo Economicus and Homo Politicus: Interpretation and Aggregation of Environmental Values." *Journal of Economic Behavior & Organization* 42 (2000): 305-322.

Nyborg, Karine. 2012. *The Ethics and Politics of Environmental Cost-Benefit Analysis*. Routledge.

Posner, Eric A., and Cass R. Sunstein. 2017. "Moral Commitments in Cost-Benefit Analysis." *Virginia Law Review* 103.

Rateman, Ty. 2007. "On the Role of Preferences and Values in Public Decisions." *Social Theory and Practice* 33, no. 2: 251-276.

Richardson, Jeff, Angelo Iezzi, Kompal Sinha, Munir A. Khan, and John Mckie. 2014. "An Instrument For Measuring The Social Willingness To Pay For Health State Improvement." *Health Economics* 23, no. 7: 792-805.

Richardson, J. and Schlander, M. 2019. "Health Technology Assessment (HTA) and Economic Evaluation: Efficiency or Fairness First." *Journal of Market Access & Health Policy*, 7(1).

Sagoff, Mark. 1981. "At the Shrine of Our Lady of Fatima or Why Political Questions are Not All Economic." *Arizona Law Review* 23: 1283.

Ubel, Peter A., Jeff Richardson, and Paul Menzel. 2000. "Societal Value, the Person Trade-Off, and the Dilemma of Whose Values to Measure for Cost-Effectiveness Analysis." *Health Economics* 9, no. 2: 127-136.

Zamir, Eyal, and Barak Medina. 2019. "Deontological Morality and Economic Analysis of Law." In White, Mark D., ed. *The Oxford Handbook of Ethics and Economics*. Oxford University Press.

Zerbe, Richard O. 2002. *Economic Efficiency in Law and Economics*. Edward Elgar Publishing, 2002.

Zerbe Richard O. 2004. Should Moral Sentiments be Incorporated into Benefit-Cost Analysis? An Example of Long-Term Discounting." *Policy Sciences* 37, 305-318.