Journal of Benefit-Cost Analysis (2023), 14: 1, 68–92
doi:10.1017/bca.2023.5

CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Distributional weighting and welfare/equity tradeoffs: a new approach

Daniel J. Acland[1] and David H. Greenberg[2]

[1]University of California, Berkeley.
[2]University of Maryland, Baltimore County.
Corresponding author: Daniel J. Acland; Email: acland@berkeley.edu

Keywords: Distributional weighting, diminishing marginal utility of income, Equity

JEL classification: D3, D6

Abstract
There are increasing calls for concrete suggestions on how to account for distributional impacts in policy analysis. Within the context of benefit-cost analysis, *per se,* one possibility is to apply "distributional weights," to inflate costs and benefits experienced by poor or disadvantaged groups. We distinguish between "utility-weights," intended to correct for the bias in willingness to pay caused by diminishing marginal utility of income, and "equity-weights," intended to account for the possibility that decision makers might have disproportional concern about the welfare of the poor or other disadvantaged groups. We argue that utility-weights are appropriate and necessary to maintain the legitimacy of BCA as a measure of aggregate welfare, but that equity-weights are inappropriate because they involve moral judgments that should remain in the domain of democratically accountable decision makers, and because they conflate information about both the welfare and equity impacts of policies, making it impossible for decision-makers to apply their own moral values to the assessment of tradeoffs between welfare and equity. We offer concrete suggestions regarding the application of utility-weights and the calculation of a set of metrics to provide intuitively comprehensible and useful information about, and allow decision makers to quantitatively assess the tradeoffs between, welfare and equity caused by specific policies.

Section 1: Introduction

Existing executive orders and Circular A-4 on regulatory analysis (2003, especially pp. 14 and 18), the official guidance document for regulatory review in the federal government, call for analysis of distributional impacts; but no concrete guidance has been provided. On the first day of his presidency, Joe Biden issued a memorandum entitled "Modernizing Regulatory Review." The memorandum directs the Director of the Office of Management and Budget to "propose procedures that take into account [among other things] the distributional consequences of regulations, including as part of any quantitative or qualitative analysis of the costs and benefits of regulations," and to

© The Author(s), 2023. Published by Cambridge University Press on behalf of the Society for Benefit-Cost Analysis. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

"provide concrete suggestions on how the regulatory review process can promote [among other things] equity." The purpose of this paper is to provide recommendations for the kinds of procedures needed to help meet the goals of the memorandum. While the Biden memorandum is directed only towards the federal regulatory impact assessment process, our recommendations are intended to improve the quality of public decision making beyond the realm of federal regulatory impact assessment.

One often proposed approach to accounting for distributional impacts is incorporating "distributional weights" into benefit-cost analysis (BCA) by multiplying the monetized value of impacts on groups at different income levels by a set of weights that vary with income, so that the willingness to pay (WTP) of those with lower income is inflated relative to that of higher-income groups.[1] The idea of weighting benefits and costs accruing at different levels of income has been discussed by scholars for many years, with the pace of publication accelerating in the 21st Century.[2] Distributional weights have been proposed as a way to address two separate and distinct issues. The first is the fact that the diminishing marginal utility of income makes BCA, as a measure of welfare, biased against the poor. In other words, because each dollar of consumption generates more welfare for a poor than wealthy person, a given change in the welfare of the poor will be represented by a smaller number of dollars in BCA than the same change in the welfare of the wealthy.[3] Thus, in terms of the distribution of utility, or welfare, BCA actually promotes *inequality*, the opposite of what the Biden memorandum calls for, and surely the opposite of what most scholars and practitioners of BCA would want. If we knew *how much* more each dollar matters to low-income individuals than to higher-income individuals, in quantitative terms, then we could adjust the willingness-to-pay for any given cost or benefit individuals receive as a result of the policy so that it represents the same change in utility or welfare, regardless of income.

The second reason one might apply weights to BCA by income is that decision makers may simply have greater concern about how policies affect the welfare of low-income groups than they do about the welfare of higher-income groups, what Adler 2013) refers to as "diminishing moral value of welfare." A dramatically unequal distribution of welfare, and the concomitant poverty that we observe in some ostensibly high-income countries, is considered by many to be unjust, unfair, or simply obscene.[4] The aggregate net benefit of a policy measured by BCA, even after adjusting for the diminishing marginal utility of income, takes no account of any concerns decision makers may have about the distribution of *welfare* impacts. Overweighting the welfare of low-income groups and underweighting that of higher-income groups, above and beyond the kind of "bias-correcting" weighting mentioned above, is seen as a way to overcome this problem, turning BCA into a tool that

[1] Our discussion is largely restricted to distribution of impacts across income groups, but parts of our approach can be applied to distribution across racial groups or other distributional categorizations.

[2] Early attempts to develop distributional weights include Eckstein (1961), Weisbrod (1968), and Feldstein (1974). A few of the more recent attempts are Layard, Nickell, and Mayraz (2008); Farrow (2021); Scarborough and Bennett (2012); and Yitzhaki (2003).

[3] Feldstein (1974) and Layard, Nickell, and Mayraz (2008), among others, presume that distributional weights would be based on this argument.

[4] For example, Graham (2008) suggests that unless a regulation designed to save lives "is neutral or yields a net gain for the poor as a group, it should not be promulgated, regardless of its consequences for society as a whole" (p. 519). Also, see Brent, 2017, Chapter 6.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

simultaneously accounts for aggregate welfare impacts and impacts on the distribution of welfare.[5]

A central thesis of this paper, and an important part of its contribution, is that these two reasons for applying weights to costs and benefits in BCA must be understood as separate, and, consequently, they should be treated very differently. In particular, we recommend that the issue of diminishing marginal utility of income should be addressed using weights, but that concerns about the distribution of welfare should not.[6] These two recommendations lead inexorably from a single principle, which is that decisions that involve both welfare impacts and distributional impacts should be evaluated using Multi-Goal Analysis. We use this term in the sense in which it is defined and explicated in Weimer and Vining (2017) and Bardach and Patashnik (2019), and extensively reviewed in Belton and Stewart (2002). The process consists of identifying the various social goals affected by a given set of policies, sometimes referred to as "decision criteria." Then the effect of each policy on each of these criteria is predicted and described (either quantitatively or qualitatively). Finally, the tradeoffs among goals, or criteria, are investigated and described.

One of the fundamental tenets of multi-goal analysis is that it must empower decision makers to assess tradeoffs among criteria according to their own value system. Analysts may go so far as to make policy recommendations on the basis of their own assessment of the tradeoffs, but a crucial feature of multi-goal analysis is that when analysts do make their own assessment of tradeoffs, the values implicit in their assessment must be made explicit, so that decision makers can apply their own judgment. (Weimer and Vining 2017, page 287.) The ability of each decision maker to apply their own value system to public decisions, and to lay bare to others the value judgments they have made, is essential to the functioning of what is referred to as "deliberative democracy."[7] Crucial to the achievement of this goal is that information about each of the criteria must be presented separately, rather than being conflated into a single index, so that decision makers can apply their own values to the assessment of tradeoffs. (Boardman et al, 2018, page 45; Vining and Boardman, 2006.)

[5] The distinction between these two reasons for weighting has not always been fully acknowledged, although in more recent publications there is increasing clarity about the distinction. In an early paper by Arnold Harberger (1978) titled "On the use of distributional weights in social BCA," the discussion is entirely about correcting for diminishing marginal utility, with no recognition that this does not actually address true distributional concerns. Adler (2008) spends five pages dissecting distributionally weighted BCA without noting the distinction, instead describing weighted BCA exclusively as a way to address distributional concerns. Then, in a later paper entirely devoted to distributionally weighted BCA, Adler (2013) is very clear about the distinction between the two issues; but ultimately settles on a set of weights that addresses both issues at once, thereby rendering the distinction impossible for decision makers to discern. Hammit (2021) writes about both equity considerations and the problem of bias caused by diminishing marginal utility of income without making any clear distinction between the two. Nurmi and Ahtiainen (2018) are clear on the distinction but claim that it is better to address the two issues simultaneously with a single set of weights. Furthermore, we note that there could be reasons to apply weights to BCA other than addressing issues of income distribution, such as correcting for differences between the valuations of health states by healthy versus unhealthy individuals. By focusing on concerns arising from the distribution of income, we do not mean to suggest that no other reasons for weighting should be considered.

[6] We realize that there could be other reasons for weighting such as the fact that any given increment in health-related quality of life may be of greater value to someone currently in poor health than to someone in good health. We do not address these forms of weighting but recognize that they could be important.

[7] This concept implies that, for a democratic decision to be legitimate, decision making requires authentic deliberation, not just the aggregation of preferences that occurs in voting (see Landemore, 2017).

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Welfare and distributional impacts are separate decision-making criteria, or goals, that are affected by many, if not most policies, and as such, decision makers need the most accurate and unbiased information possible about each of these criteria separately. Circular A-4, the primary guidance document for regulatory review in federal agencies of the U.S. government, makes it clear that multi-goal analysis is the preferred approach to addressing tradeoffs between welfare and distributional impacts, stating that regulatory analysis "should provide a *separate* description of distributional effects…so that decision makers can properly consider them *along with* the effects on economic efficiency." (OMB 2003, emphasis added.) Weighting to account for diminishing marginal utility of income generates unbiased information about aggregate welfare. Further weighting, to account for moral or ethical concerns about the distribution of welfare, conflates information about welfare with information about the distribution of welfare, contrary to the tenets of multi-goal analysis.[8] Thus, our recommendation boils down to this: willingness-to-pay should be weighted to account for diminishing marginal utility of income, but no further weighting should be done. These recommendations are not, strictly, matters of opinion. They are the logically inevitable conclusions that lead from the premise that multi-goal analysis is the appropriate form of analysis when multiple social goals are affected by a policy.

However, having recommended that weighting to account for moral or ethical concerns about the equitability of the distribution of welfare should be avoided, we provide a set of metrics that allow for quantitative assessment of the tradeoffs between the impacts of a policy on welfare, on the one hand, and the distribution of welfare, on the other hand, and also describe a stepwise procedure for assessing tradeoffs in the simplest manner possible. As a package, these recommendations, metrics, and procedures maximize decision makers' ability to assess tradeoffs between the distinct criteria of welfare and distribution.

In section 2, we discuss weighting to address bias in willingness-to-pay (WTP) as a measure of welfare, which we refer to as utility-weighting. In section 3, we discuss additional weighting to address concerns about the distribution of welfare, which we refer to as equity-weighting. In section 4, we propose and discuss our metrics for assessing the tradeoffs between welfare impacts and equity concerns. In section 5, we present concrete recommendations and apply them to a set of real-world cases. Section 6 concludes.

Section 2: Utility-weighting

2.1. Justification for Utility-Weighting

Aggregate welfare is understood by normative ethicists to have ethical or moral weight in decision making and to be a value that is separate from other values such as equity and rights (Kagan, 2018). As such, it is a social goal about which decision makers need information, in multi-goal analysis, that is not biased by the diminishing marginal utility of income. We use the term "utility weights" to refer to weights that are applied to different income groups to achieve the necessary debiasing. When utility weights are applied to BCA, it becomes an

[8] We show below that weighting for moral concerns quite literally makes it impossible for a decision maker to discern even whether either of the goals has been advanced or retarded. An increase in the metric that results from weighting for moral concerns could be caused by an increase in welfare accompanied by a decrease in distributional equality, by an increase in equality accompanied by a decrease in welfare, or by an increase in both.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

unbiased measure of aggregate welfare, and, as referenced in the introduction, unbiased information about social goals is one of the necessary ingredients in multi-goal analysis.

Unweighted BCA is often referred to as a measure of "allocative efficiency," which Boardman et al. (2018) describe as "a situation in which resources, such as land, labor and capital, are deployed in their highest valued uses in terms of the goods and services they create." It is clear that the term "valued" means valued in unweighted dollars. The sense in which this situation is maximally efficient is not addressed, as it does not actually involve maximizing what actually matters to the lived experience of individuals, which is welfare. Regardless of how the term "allocative efficiency" is defined, BCA simply is not a measure of what is best for society in terms of welfare. Nonetheless, the interpretation of BCA as a measure of aggregate welfare has been embraced by many scholars. Adler and Posner (2006) in their attempt to provide "New Foundations" for BCA, endorse the welfare interpretation explicitly, defining BCA as a measure of "restricted preferentialist welfare," while Sunstein (2018), throughout his treatise on the "Cost-Benefit Revolution," consistently uses the term "welfare" to describe the value measured by unweighted BCA. Harberger (1978), one of the first scholars to discuss what we are referring to as utility weights, also used the term "welfare," without comment, to describe the value measured by WTP.[9] In addition, a case can be made that the acceptance of the interpretation of BCA as a measure of aggregate welfare is implicit in the universally accepted term "net benefit." The difference between the WTP of the winners and the willingness-to-accept (WTA) of the losers is, more or less without exception, referred to as "net benefit" throughout the field of BCA – a term that is commonly implicitly used as a synonym for welfare (if perhaps unconsciously). It is perhaps understandable that scholars and practitioners of BCA explicitly or implicitly embrace the aggregate welfare interpretation of unweighted BCA since, as we have stated, aggregate welfare is the thing that normative ethicists and moral philosophers believe decision makers ought to be informed about.

Some have argued that unweighted BCA is an imperfect, or "rough and ready" proxy for welfare (Adler and Posner 2006), but if the proxy is systematically biased against one group, then it is not just imperfect, it is unacceptable. For example, a policy that generates a benefit of $100 for a corporate attorney while imposing a cost of $90 on a custodial worker is an improvement in allocative efficiency, but to say that $10 is a rough-and-ready approximation of the welfare benefit to society will strike many as morally repugnant. Others have argued that the consistent application of unweighted net benefit to policy making will result in all groups benefiting in the long run, as some policies will benefit the wealthy, and others will benefit the poor, so that on net there will be an improvement to both groups. However, even in the long run, unweighted net benefit will result in greater benefits to the wealthy than to the poor, because it will consistently preference policies that benefit the wealthy over those that benefit the poor. For example, if two policies generate a welfare benefit of X and a welfare cost of Y, but the first assigns the benefit to the wealthy and the cost to the poor, and vice versa for the other, unweighted BCA will rank the former above the latter, even if both pass an unweighted net-benefit test. To determine whether public decision-making increases aggregate welfare, policy-by-policy in the short-run, or on net over the long-run, we must apply utility weights.

[9] Harberger (1978) repudiated distributional weighting because it reduces allocative efficiency. He made no mention of the fact that he was defining efficiency in terms of dollar valuation instead of actual welfare. What he viewed as being lost was not, in fact, welfare.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

A fundamental challenge must be addressed before discussing the practicalities of utility-weighting, which is the problem of interpersonal welfare comparisons: no matter what metric one uses in the attempt to capture welfare, the same number may refer to different levels of welfare across individuals. Even if two individuals have identical incomes and identical preferences, the fact that they have identical willingness-to-pay for a good that increases their welfare does not necessarily mean that they get the same welfare benefit from that good. One may simply get more welfare out of the good than the other – holding their preferences and income constant – perhaps because of their temperament, upbringing, or experience. Consider two individuals, with identical preferences for fly fishing versus all other goods, identical income, identical life circumstances, and identical WTP for a fly-fishing trip. Even if they have identical experiences on a fly-fishing trip, one may simply have a better time, perhaps because he or she is simply inherently easier to please. Alternative metrics do not solve the problem. For example, when researchers attempt to measure welfare with survey questions that ask respondents to rate their "subjective wellbeing" on a scale of, let's say, one to ten, there is no way for them to know whether a response of six out of ten represents the same level of welfare from one individual to another, and not only because the scales are typically ordinal. Even if the scale could somehow be thought of as cardinal, one individual might simply have a higher reference point for wellbeing, so that a six out of ten refered to a higher level of utility for them than for another individual. The impossibility of an interpersonally comparable welfare metric has been recognized since the beginning of the ordinalist revolution pioneered by Fisher and Pareto (see Lenfant., 2012).

The problem of interpersonal comparisons in the measurement of aggregate welfare led the ordinalists to introduce the potential Pareto criterion (PPC) as a way to sidestep the problem. Rather than wrestling with whether or not a policy increases aggregate welfare, they preferred to assess the desirability of a policy by determining whether it is a "potential Pareto improvement." A policy is a potential Pareto improvement to society if those who gain from the policy could potentially transfer enough money to those who lose to make the losers better off, or at least no worse off, and still come out ahead. If we embrace the PPC and interpret BCA as an implementation of it, then BCA measures the difference between aggregate willingness to pay (WTP) for the gains and aggregate willingness to accept (WTA) for the losses, which could be thought of as the magnitude of the improvement. Interpreting BCA as an implementation of the PPC puts it on solid ground with respect to interpersonal welfare comparisons: it is simply no longer an attempt to measure aggregate welfare.

If this is how we understand BCA, then clearly we must not apply utility weights, because if we do, we will no longer have accurate measures of WTP and WTA. Consider a policy that has been determined to be a potential Pareto improvement: WTP is greater than WTA. If those who gain have higher income than those who lose, utility-weighting would inflate WTA relative to WTP, so that the difference between utility-weighted WTP and utility-weighted WTA would be smaller than without the weights and might indeed be negative. Utility-weighted BCA cannot be counted on to determine whether the winners could compensate the losers in monetary terms. Thus, if we interpret BCA as an implementation of the PPC, we must not apply utility weights, we must accept that BCA does not measure aggregate welfare, and we must accept that it cannot reliably be used by decision makers to assess the tradeoffs between welfare and distributional impacts in multi-goal analysis. If one accepts our premise that multi-goal analysis is the appropriate way to empower decision makers to assess tradeoffs between welfare and distributional impacts, this concession is

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

anathema. If, instead, we apply utility weights and interpret BCA as a measure of aggregate welfare, we must confront the interpersonal comparisons problem.

We do not have a solution to the problem of interpersonal welfare comparisons between specific individuals, but we do not need one because we are only concerned with comparing benefits and costs between *groups.* All that this requires is that the distribution of different types of individuals, with respect to the welfare or utility they get out of goods for which they have equal WTP, be the same in the different groups being compared. For example, suppose that the costs of a policy accrue to consumers of a particular good who are primarily high-income, while the benefits are received by the relatively low-income members of a particular occupation. A problem would arise if either the consumers or the workers were disproportionately individuals for whom a given WTP represents relatively little welfare. Utility-weighted BCA would be favorably biased towards that group because a given dollar value of the costs or benefits to that group would represent less welfare than the same aggregate WTP among the other group.

We are not aware of any empirical evidence of this kind of disproportionality, and indeed, the very fact of noncomparability makes it impossible to gather any such evidence, as it would necessarily require some measure of welfare, which, in turn, would be subject to the problem of interpersonal comparisons. Nonetheless, to justify interpreting BCA as a measure of aggregate welfare and applying utility weights, we must make the assumption that there is no disproportionality of types across groups. The rapidly growing literature on the measurement of happiness or subjective well-being (SWB) also makes this assumption (Wiener et al, 2013; Frey et al, 2010). Indeed, the case could be made that those within the field of BCA who think of BCA as a measure of welfare are implicitly making the same assumption. Harberger made a poignant and relevant observation in 1978: "[A]lthough most economists have come to shun interpersonal comparisons of utility whenever they feel they can avoid or circumvent them, they usually make them when they have to." (Harberger 1978.) Assuming that the distribution of types of individuals is the same across groups is considerably more innocuous than assuming interpersonal comparability at the individual level and would appear to be a modest compromise to make in pursuit of a measure of aggregate welfare that is not unethically biased against the poor.

2.2. Constructing Utility Weights

How then should utility weights be constructed? One approach would be to convert dollars into "utils" by multiplying the WTP of an individual of a given income level by their marginal utility of income at that income level. The number of utils assigned to any given quantum of welfare would be the same, regardless of the income level of the individual. The problem with this approach is that the utility of income is invariant to any affine transformation. In other words, if $U(y)$ represents some preferences over income, then for any constants a and b, $U'(y) = a + b*U(y)$ will represent the same preferences (provided that b is positive). But because the two functions have different first derivatives, they generate different marginal utilities and thus different measures of weighted net benefits.

The standard solution to this problem in the literature is, rather than using the marginal utility of each individual at their current income level, to instead use the ratio of their marginal utility at their current income to the marginal utility they would have at some reference income, such as the median income in the population (Cowell and Gardiner, 1999; Adler, 2013; Nurmi and Ahtiainen, 2018). This ratio is the same across all affine

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

transformations of the utility function, because b is in both the numerator and the denominator (while a drops out when taking derivatives). This is true both for a given individual at two different levels of income and for two individuals with the same preferences at different levels of income, provided that we use the same transformation of the utility function to represent the preferences of the two individuals.

This ratio measures the increase in utility generated by an additional dollar at y_0, as a multiple of the increase in utility generated by an additional dollar at the medium income level, y_m. For example, if at y_0, the marginal utility of income is 4, while at y_m it is 2, then the utility weight on WTP at y_0 would be $4/2 = 2$. Gaining a dollar when one's income is y_0 generates twice as much utility as gaining a dollar when one's income is y_m. Inflating dollars gained when income is y_0 by a factor of 2 generates an unbiased metric of utility. A utility-weighted dollar now represents the same increase in utility regardless of income level. Consequently, we can think of the utility-weighted WTP of an individual at income y_0 for some quantum of welfare as the WTP they would express for that quantum of welfare if they had median income. The bias has been removed, and the result has an intuitive comprehensible interpretation: the utility-weighted net benefit to society of a policy can be thought of as the aggregate WTP one would observe in a hypothetical society in which everyone had income equal to the median.[10]

2.3. Computing Utility Weights

Computing utility weights requires that we have some way to estimate the ratio of marginal utilities between income levels. There is a large literature that attempts to do this.[11] In order to make the estimation process feasible, it is necessary to make an assumption about the functional form of utility. In the methodologies we rely on below, the standard assumption is that preferences can be represented by an isoelastic utility function:

$$U(y) = \frac{y^{1-\varepsilon} - 1}{1-\varepsilon} \text{ if } \varepsilon \neq 1 \text{ and } U(y) = \ln\,(y) \text{ if } \varepsilon = 1 \tag{1}$$

where y is income and ε is the elasticity of marginal utility with respect to income.[12] The isoelastic function has the desirable property of constant elasticity.

[10] From the perspective of this interpretation, the per-capita utility-weighted net benefit would appear to be comparable to the "equally distributed equivalent" (EDE), which is the level of any given individual-level measure (in our case WTP) that, if every individual in the population had that measure, would generate the same value of whatever social welfare function is being applied (See Sheffiff and Maguire, 2020 and Fleurbay, 2010). In our case, this would be the level of WTP for the impacts of a policy that, if everyone in society experienced it, would generate the same utility-weighted net benefit as the actual impacts of the policy. This similarity is illusory, however, because, even if everyone experienced the same WTP for the impacts of a policy, they would still have different incomes, and thus different utility weights. For example, consider a society with two individuals with different incomes, one with a utility weight of 2 and the other with a weight of .5, and suppose a policy generated a benefit to the individuals for which they were willing to pay $10 and $40, respectively. The utility-weighted net benefit is $40, which is $20 on a per-capita basis. If both individuals were given a welfare impact for which they had WTP of $20, the utility-weighted net benefit would be $50, so the per-capita utility-weighted net benefit is not the same thing as the EDE. This is because, even if they experience the same WTP for the impacts of the policy, their incomes, and thus their utility weights, remain different.

[11] A large number of references to this literature can be found in Groom and Maddison, 2019; Acland and Greenberg, 2023; and Havranek et al., 2013.

[12] It is also, not coincidentally, the Arrow-Pratt coefficient of relative risk aversion.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Given this utility function, the marginal utility of income is $\frac{dU}{dy} = \frac{1}{y^{\varepsilon}}$. If we normalize the weight on individuals of median income to be 1, and if we let y_i denote the income of group i, and y_m denote median income, then the appropriate utility weight to place on the WTP of group i is

$$w_i = \frac{dU_i}{dy_i} \bigg/ \frac{dU_m}{dy_m} = \left(\frac{y_m}{y_i}\right)^{\varepsilon} \tag{2}$$

The information necessary to compute y_m/y_i for different income groups is obtainable in many countries. In the U.S., it can be found in federal government statistics such as the U.S. Census Bureau's Annual ASEC Survey.[13] In addition, computing utility weights for various income groups requires an estimate of ε. In Acland and Greenberg (2023), we conduct a meta-analysis of 158 studies, comprising 1,727 estimates of ε, that use a variety of methodologies. The conclusion from our meta-analysis is that the mean estimate of ε in the literature is 1.61 with a 95% confidence interval of 1.18 to 2.04. To give a sense of what these numbers mean, Table 1 shows the utility weights implied by our mean estimate and the upper and lower bounds of our confidence interval for several percentiles of the distribution of income in the U.S. in 2021. Because income is higher in larger households than in smaller households, and the average size of U.S. households is 2.5, the income amounts used to compute the utility weights are a weighted average of the incomes of two-person and three-person households. As shown in Table 1, the weights must equal 1 at the 50th percentile because at that percentile $y_m = y_i$ and hence $y_m/y_i = 1$.

This table contradicts two common practical arguments against utility-weighting. The first is that under plausible conditions, utility-weighting generates weights that seem intuitively implausible. Banzhaf (2011) points out that if $\varepsilon = 2$, the weight on a person with income of $10,000 is 100 times the weight on a person with income of $100,000. What Banzhaf does not mention is that an income of $10,000, in most parts of the U.S., puts an individual in conditions of such extreme and abject poverty that the welfare impact of even quite small sums of money may be very large. For someone at that level of poverty, an extra $100 a month might mean the difference between two meals a day and three, whereas for

Table 1. *Utility weights at selected income percentiles for average household size.*

		Elasticity Estimate (ε)		
Percentile	Household Income, 2021	1.18	1.61	2.04
10th	$23,250	4.20	7.08	11.94
25th	$44,250	1.96	2.51	3.21
30th	$50,900	1.66	2.00	2.41
50th	$78,400	1.00	1.00	1.00
75th	$128,000	0.56	0.45	0.37
90th	$199,365	0.33	0.22	0.15
95th	$272,000	0.23	0.13	0.08
99th	$554,500	0.10	0.04	0.02

[13] We recognize that determining the income level of the different groups impacted by any given policy may be non-trivial. We discuss this later.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

someone at \$100,000, the same sum is more likely to mean one nice evening out each month. A difference in welfare impact of 100 to 1 might not be so unreasonable. Furthermore, at less extreme differences in income, the relative weights are less extreme. Considering the 25th and 75th percentiles (\$44,250 and \$128,000 respectively), even with $\varepsilon = 2$, the weight on the lower-income household is only approximately 8 times the weight on the higher-income household, and at the 1.61 value of ε we estimate in our meta-analysis, it is approximately 5.5 times higher. The weight on the 10th percentile is 32 times that on the 90th percentile, but again, these are households with extreme differences in standard of living. Careful consideration of the lived experience of households at these levels of income might suggest that these numbers are not unwarranted.

The second practical argument against utility-weighting is that there is little hope of consensus on the appropriate set of weights (see Mishan, 1988, pp. 201, 210). In fact, as our meta-analysis shows, the extensive literature estimating the elasticity of marginal utility of income converges on a relatively narrow range, and as the literature on distributional weighting, without exception that we know of, uses an isoelastic utility function to derive weights, this suggests that there is, in fact, a strong basis for a consensus on the appropriate utility weights.

Some have argued that the selection of distributional weights involves subjective value judgments and thus should not be the responsibility of analysts (see Dasgupta, Marglin, and Sen, 1972; Little and Mirrlees, 1974; and Ray, 1984). This is not the case when it comes to utility-weights (although, as discussed in the following section, it may be true in addressing equity issues). The relative value of increases in income at different income levels – i.e., the marginal utility of income – results from the preferences of individuals, and the observation of preferences is a technical matter. Furthermore, it is something that can be estimated using standard techniques, as we discussed above. As such, the determination of utility weights is the responsibility of economists and analysts, not decision makers. Alfred Marshall rendered his opinion on this in 1885, saying "taking account of the fact that the same sum of money measures a greater pleasure for the poor than for the rich [is a] task [that] most properly belongs to the economic organ" (Marshall 1885). This view is essential to our recommendations. The bias caused by diminishing marginal utility of income is a technical problem in the measurement of welfare and can be empirically addressed by observing behavior that reveals individual preferences. It does not require moral or ethical judgments. As such, it lies squarely in the domain of the analyst.

Section 3: Accounting for Equity Impacts

3.1. Equity-weighting

Once BCA has been converted into an unbiased measure of aggregate welfare by the application of utility weights, it remains the *unweighted* sum of the *welfare* impacts on different groups. In other words, it places the same weight on the *welfare* of individuals at different levels of income, and thus does not account for the possibility that decision makers may simply have greater concern for the welfare of individuals at the lower end of the income distribution. Utility-weighting merely overcomes a source of bias in the measurement of welfare. Adler (2013) and Nurmi and Ahtiainen (2018) both call for using weights to address what the former refers to as diminishing "marginal moral value of utility." In other words, the

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

higher a person's existing welfare, the less society should "care" about increases in their welfare. We will refer to this kind of weighting as "equity-weighting."

Adler's terminology clearly implies that this kind of weighting involves moral value judgments. An issue of fundamental importance is who should make these kinds of judgments. This is where the biggest difference lies between utility-weighting and equity-weighting. As we suggest in section 2, decisions about what utility weights to use – being primarily a matter of observing individual preferences – should lie in the hands of economists and analysts. By contrast, decisions about how to assess tradeoffs between welfare and equity, which is to say decisions about how much concern society should have about the distribution of welfare across income groups, relative to how much concern society should have for aggregate welfare, belong in the hands of decision makers. Adler (2013) comes close to taking this position:

> "The use of distributional weights does raise questions of institutional role. An unelected bureaucrat might feel that it would be legally problematic, or democratically illegitimate, for her to specify weights. Who in government gets to act on contestable moral preferences is a complicated (and itself contestable) question of law and democratic theory."[14]

For us, however, the question is not who should choose the weights when doing equity-weighting. The question is whether equity-weighting *per se* helps or hinders decision makers' ability to make their own value judgments about welfare/equity tradeoffs in the first place. Recall the relevant desideratum for multi-goal analysis, which provides information about the effect of a policy on each of society's disparate goals. If presented separately, this information should allow decision makers to directly observe policy effects on each goal and apply their own values.

By combining information about welfare and information about distribution into a single metric, equity-weighting attempts to do precisely the opposite. The problem is quite fundamental: this metric provides no information about welfare and no information about distribution. The fact that a certain policy is assigned a certain number by an equity-weighted BCA tells us nothing about how much welfare it generates and nothing about how it affects equity. The same number could be achieved by a policy that generates a great deal of welfare for relatively wealthy people while imposing a significant cost on the poor or by a policy that closes the gap between the rich and poor while actually lowering aggregate welfare. For example, the first policy might generate a utility-weighted benefit to the wealthy of 100 and a cost to the poor of 25, a gain in aggregate welfare but an increase in the welfare gap. The second policy might generate a utility-weighted cost to the wealthy of 100 and a benefit to the poor of 75, a loss of aggregate welfare, but a decrease in the welfare gap. With an equity weight of 2 on the welfare of the poor, both policies generate an equity-weighted net benefit of 50. A decision maker learns literally nothing about the separate welfare and equity impacts of either policy from such equity-weighting. It is impossible for him or her to assess the

[14] One could argue that the important distinction is not whether the bureaucrat is elected or unelected, but whether he or she has in some way been tasked with making decision in the best interest of society as a whole, whether by dint of being appointed by an elected official, or by being charged by some statute or legal ruling to do so. In our view, any such decision maker might consider themselves justified in specifying equit weights, but someone acting in the role of analyst should not.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

tradeoffs, in just the same way that knowing the volumes of two boxes provides no information whatsoever with which to assess tradeoffs between depth, width and height. Applying equity weights is antithetical to the principles of multi-goal analysis.

Two claims could be made in response to this critique of equity weights. First, one could argue that if a policy imposes costs on the poor and has positive utility-weighted net benefits, and if the application of an appropriately chosen equity-weight changes the sign of net benefits from positive to negative, it is evidence that the equity impact of the policy outweighs the welfare impact. This, of course, assumes that the chosen equity weights align with the ethical or moral values of the decision maker reading the analysis, which brings us back to the goal of empowering decision makers to apply their own values. If the analyst chooses the equity weights, the decision maker is no longer empowered to apply their own values. To overcome this problem, and empower decision makers to apply their own values, the analyst could do a sensitivity analysis, providing equity-weighted net benefits for a range of different weights, thereby allowing decision makers to choose their preferred weight.

However, if the goal is to allow decision makers to determine whether distributional impacts outweigh welfare impacts, then there is arguably a superior way to sensitivity analysis to do so, which is to compute the "breakeven" equity-weight, that is the weight that causes net benefits to change sign. The advantage of breakeven equity weights, as discussed in more detail later, is that they present decision makers with a discreet, binary assessment: "is my preferred weight above or below this threshold?" This decision is arguably much easier to make than the continuous assessment: "which of these is closest to my preferred weight?" It may be quite hard for decision makers to translate their own value judgment into a specific preferred equity weight; but in many cases, it may be relatively easy for them to assess whether their value judgment translates into a weight above or below the breakeven threshold.

Furthermore, applying equity weights to a policy that generates welfare/equity tradeoffs to see if the sign of net benefits changes only works when comparing one policy to the status quo. It is tempting to hope that if the equity-weighted net benefit of one policy is greater than the equity-weighted net benefit of another, it must be the case that the first policy involves a more favorable tradeoff between distribution and welfare; but this is not the case. As stated earlier, the fact that one policy has a higher equity-weighted net benefit than another tells us nothing about either the welfare comparison or the distributional comparison between the two policies. What is called for when comparing two policies is information about the tradeoff between welfare and distribution implied by moving *from one to the other*, not information about how each policy compares to the status quo. And as shown in Appendix 1, it is not possible to apply equity weights (or even the breakeven weight) in a consistent and defensible manner when making this comparison.

The second claim that can be made in response to our critique of equity-weighting is that when equity weights are applied to utility-weighted net benefits, the resulting metric is a measure of a unitary moral value, of which welfare and distribution are merely instrumental parts. In other words, welfare and distributional impacts are not separately, and in and of themselves, relevant to the assessments of decision makers. The unitary moral value in question is referred to as "social welfare."[15] In particular, the typical recommendation in the

[15] It is true that utility-weighted BCA is a form of "social welfare function," in particular a utilitarian social welfare function; but we are going to use the term to refer to metrics that combine information about welfare and distribution, which the utilitarian social welfare function does not.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

social-welfare measurement literature is that equity weights be derived from an iso-elastic function, parallel to the one we use in the determination of utility weights, thereby rendering equity-weighted BCA a way to implement a "continuous prioritarian" social welfare function. It gives "priority" to the welfare of the poor in a "continuous" way, which is to say that for any given increment in welfare, there is an incremental decrease in equity weight (Adler, 2019). The resulting unitary value is sometimes referred to as the "overall wellbeing" of society or simply the "good" of society (Kagan 2018). Against those who consider the abandonment of separate measures of welfare and distribution in favor of this unitary metric of what is good for society, we have no defense. Their position is antithetical to our view that welfare and distribution are intrinsic values that have moral weight in and of themselves, and about which decision makers need separate information so that they can assess the acceptability of the welfare/equity tradeoffs created by a policy.

3.2. Metrics for assessing welfare and equity impacts

In this section, we propose a pair of metrics that can be used for quantitatively assessing the tradeoffs between the welfare and equity impacts of a policy relative to the status quo. (In Appendix 1 we demonstrate how the metrics can be applied to the comparison of two or more non-status-quo policies.) There are many strategies for quantification of distributional impacts. Farrow (2011) suggests, for example, tables of disaggregated impacts by group, frequency plots, Lorenz curves, tables of quantiles, and a number of scalar measures such as variance, Gini coefficient, and the Atkinson index. While it is not our purpose to critique these strategies, we note that none of them provide information about the *tradeoff* between distributional impact and welfare impact. They simply tell us something about what the distributional impact *is.* We are not proposing any alternative metric for quantifying distributional impacts. Nor are we saying that quantitative measurement is superior, in any given instance, to qualitative analysis. Rather, our proposed metrics are intended to be used to quantitatively assess the acceptability of a given *tradeoff* between welfare and distributional impact. Furthermore, the second metric we propose can be used to assess tradeoffs between welfare and any univariate metric of distributional impact. We recommend using these metrics to account for distributional concerns, instead of using equity-weights.

3.2.1. Breakeven equity-weight

The first metric we consider is the breakeven equity weight, which we mentioned above. If a policy fails the utility-weighted BCA test because the welfare benefits to the poor are outweighed by the welfare costs to the non-poor, then there will be some number greater than one, which, when multiplied by the utility-weighted benefit to the poor, will cause the utility-weighted net benefit to society to rise to zero. We call that number the "breakeven equity weight." For any weight above that number, the net benefit will become positive. (The breakeven equity weight works in reverse when costs are to the poor, benefits are to the non-poor, and the policy has positive utility-weighted net benefits.)

The breakeven equity weight is the ratio of the net impact (cost or benefit) on the non-poor to the net impact (benefit or cost) on the poor. Returning, for example, to the two policies in section 3.1 that had different welfare and distributional impacts, but the same equity weighted net benefit, the first has a breakeven equity weight of $\frac{100}{25} = 4$ (because the cost

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

to the poor is one-quarter of the benefit to the non-poor), while the second policy has a breakeven weight of $\frac{100}{75} = 1.33$ (because the benefit to the poor is three-quarters of the cost to the non-poor).

We view the breakeven equity weight as a number that is intuitively comprehensible. Consider the first policy above with a breakeven equity weight of 4. One interpretation of that number is that if we deem the welfare cost to the poor to be at least four times more important to society than the welfare gain to the non-poor, then we might conclude that the equity loss is not worth the welfare gain. In the case of the second policy, if our concern for the welfare of the poor is at least a third greater than our concern for the welfare of the non-poor, then the equity gain will outweigh the welfare loss. Clearly, on the basis of this metric, if we were deciding between the first policy and the status quo, we would be less inclined to adopt that policy than if we were deciding between the second policy and the status quo, because the first policy requires much greater concern for the poor.[16]

It is crucially important to understand that this pair of breakeven equity weights does not allow us to determine whether the second policy is preferable to the first in a direct comparison (rather than a one-by-one comparison of each policy to the status quo). When comparing two non-status-quo policies, what matters is the tradeoff implied by choosing one of the policies over the other. However, it is not possible to use breakeven weights to compare two non-status-quo policies, an issue we address in Appendix 1.

As stated earlier, in many cases the breakeven weight will be relatively straightforward for decision makers to make use of. It isn't necessary for a decision maker to know specifically what their own preferred weight on the welfare of the poor is. It is only necessary for them to have a sense of whether it is above or below the breakeven weight. There will be cases in which this question is hard to answer, but in many instances the assessment is likely to be relatively straightforward because the breakeven weight will clearly be above(below) the upper(lower) bound of what they consider acceptable.

In Appendix 2 we provide a method for applying breakeven weights when there are three affected groups.

3.2.2. Welfare/equity ratios

Weimer and Vining (2017) make clear that when efficiency (read welfare) impacts have been quantified and monetized, and there is a single non-welfare social goal that can be quantified, then the correct form of analysis to use to inform decision makers about the tradeoffs between welfare and the other goal is cost-effectiveness analysis, as it can be used to determine the welfare cost of each unit of gain in the other goal, or vice-versa. Suppose the non-welfare goal is welfare equality between income groups, and the metric of equity impact is the increase or decrease in the welfare gap between two groups (let us say, loosely, a higher-income group and a lower-income group), the welfare/equity ratio would be the ratio of the utility-weighted net gain or loss of welfare and the increase or decrease in welfare gap. For a policy with positive utility-weighted net benefit that generates an increase in the welfare gap, this ratio tells us the amount of welfare gained in exchange for each dollar of increase in the welfare gap; and for a policy with negative utility-weighted net benefit that generates a decrease in the welfare gap,

[16] Though clearly these decisions would also have to consider the fact that the first policy increases welfare, while the second policy decreases welfare. Consequently, both policies present welfare/equity tradeoffs.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

this ratio tells us the amount of welfare lost in exchange for each dollar decrease in the welfare gap.[17]

This need not be the only welfare/equity ratio used to assess tradeoffs between welfare and equity. Those who are concerned about income inequality are typically also concerned about poverty in absolute terms. Thus, one could compute the ratio of the welfare gain(loss) to society as a whole to the absolute cost(benefit) to the poor, which would indicate how much aggregate welfare we gain(lose) in exchange for each dollar of negative(positive) welfare impact on the poor. Indeed, one could compute the ratio of the welfare impact of a policy and any univariate measure of equity impact, such as the Gini coefficient or the Atkinson coefficient.[18]

Note that welfare/equity ratios are *not* metrics of some unitary value that replaces welfare and distribution. Rather, they are measures of the *tradeoff* between the two. As with the breakeven weight, they allow decision makers to compare the tradeoff between welfare and distribution created by the policy with their own sense of what is acceptable. Unlike BCA, cost-effectiveness analysis does not provide a simple yes/no rule. A policy increases welfare if its utility-weighted net benefit is positive, and vice versa, but a policy can only be considered "cost effective" in relation to some threshold. In the current case, the relevant threshold is the decision maker's own value judgment.

To illustrate, consider a policy that, after utility-weighting, results in gains to the non-poor of \$200 and costs to the poor of \$50. Thus, the aggregate impact on welfare is a gain of \$150. The impact on the welfare gap (what we might call the "distributional impact") is \$250 (\$200 + \$50), while the impact on the poor (what we might call the "poverty impact") is \$50. The ratios we are proposing are thus:

- Welfare gain per unit of distributional impact = \$0.60 (\$150/\$250)
- Welfare gain per unit of poverty impact = \$3.00 (\$150/\$50)

A decision maker who is provided with these metrics might reason thus:

> "This policy has positive net benefits relative to the status quo, which is good. It is true that it imposes a cost on the poor, but that cost is relatively low, and welfare gain per dollar of cost to the poor is relatively high. Perhaps, it is worth accepting the loss on this dimension of equity in exchange for this much net benefit. On the other hand, the welfare gain per dollar of distributional impact is quite low."

The decision maker would then need to weigh the trade-offs between net benefit and the two dimensions of equity in order to make their final determination. For example, he or she might conclude that their concern about absolute poverty outweighed their concern about distribution and that, on balance, the gain in welfare was worth the equity impact.

[17] If lumping the affected population into two groups is deemed an excessive simplification, an alternative metric of distributional impact could be substituted for the welfare gap measure, such as the change in the Gini coefficient or in the Atkinson coefficient. These measures have the advantage of allowing for a continuous distribution of income levels but may be harder to compute in practice. The point is that cost-effectiveness analysis can be used to make a quantitative assessment of the tradeoff between welfare and any measure of distributional impact, provided that the tradeoff can be captured in a single measure.

[18] Furthermore, though our recommendations are explicitly with respect to income distribution, the welfare/equity ratio could be computed for measures of distribution across groups defined by race/ethnicity, gender, geography, or any other categorization within with there are deemed to be vulnerable groups.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Alternatively, consider a policy that generates a utility-weighted benefit to the non-poor of $1,000 at a utility-weighted cost to the poor of $250. The net benefit is considerably larger than is true of the previous policy, and the welfare gain per dollar of poverty impact remains at $3.00 ($750/$250), while the welfare gain per dollar of distributional impact remains at $0.60 ($750/$1,50). However, the absolute poverty impact is now quite high. Perhaps the decision maker would now feel that the relatively high welfare benefit per unit of poverty impact is unacceptable because it is outweighed by what they feel is an unacceptable absolute poverty impact. These are precisely the kinds of value judgments that democratically accountable decision makers need to be empowered to make, using precisely the kind of multi-goal analysis taught in every school of public policy. Our proposed metrics make that possible.

To place these metrics into the multi-goal analysis framework upon which our recommendations are predicated, the process begins with providing information about each of society's goals *separately*, which requires that equity weights not be applied to costs and benefits. Then decision makers are confronted with the task of assessing the tradeoffs among goals and determining whether those tradeoffs are acceptable. Our approach, using break-even equity weights and welfare/equity ratios, provides quantitative measures of the tradeoffs generated by specific policies. These quantitative measures are designed to be easily and intuitively comprehensible to decision makers.

Section 4. Implementation of utility-weights and equity metrics

We next offer the following set of recommendations for the treatment of bias caused by diminishing marginal utility of income and of equity considerations in BCA. We consider these recommendations to be necessary for BCA to be a valid measure of aggregate welfare, and thereby provide the most useful information for assessing tradeoffs between welfare and equity. After these recommendations, we offer a simplified procedure that may be applicable in many cases.

4.1. Recommendations

As stated in the introduction, the following recommendations follow from the premise that policy decisions that implicate multiple social goals should be informed by multi-goal analysis.

1. Benefit-cost analysis should be thought of as a measure of aggregate welfare, which requires that utility weights should be used so that BCA is not biased against the poor. Without this step, at a time when distributional impacts are of increasing concern in the policy-making domain, BCA is at increasing risk of being dismissed as an illegitimate form of policy analysis because it is based on principles that are widely acknowledged to be ethically and morally invalid.
2. Being primarily a technical matter of measuring individual preferences, the decision of what numerical utility weights to apply should lie in the hands of economists and analysts, not decision makers. This upholds the principle that in matters concerning individuals' pursuit of their own welfare, decision makers should defer to individuals' preferences. The most practicable method for determining utility weights is to assume an isoelastic utility function, normalize the weight at median income to 1, and compute weights for

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

other income levels using equation (2). The utility-weighted net benefit to individuals of any given income level should be interpreted as the aggregate WTP for the benefits and costs that those individuals would express if their incomes were at the median.

3. Equity weights should not be applied to BCA. The responsibility for assessing the welfare and equity impacts of a policy should lie in the hands of decision makers. Equity-weighted BCA does not contain the information decision makers need to weigh these fundamentally different types of impact because it conflates welfare and equity into a single number that entirely occludes information about welfare and equity as separate policy goals. Decision makers should be presented with information about welfare and information about equity separately, so that they can directly assess the tradeoffs between the two.
4. Our equity metrics should be computed and reported to decision makers as quantitative tools for assessing welfare/equity tradeoffs. When more than one non-status-quo policy is under consideration, the steps outlined in Appendix 1 should be used.

In addition, it is always appropriate to present utility-weighted costs and benefits disaggregated by group – however those groups may be defined – preferably in a table, with costs and benefits listed down the rows and groups listed across the columns. This allows for direct inspection of the distribution of welfare impacts and provides all of the information necessary to compute metrics of the sort we recommend. Kerry Krutilla (2005) has written an excellent explication of how this kind of table can inform decisions with distributional implications.

4.2. Steps

We offer a stepwise procedure for making use of utility-weighting and for assessing whether there is a tradeoff between welfare and distribution, in the case in which one policy is being compared to the status quo. One intent of the steps is to minimize the extent to which challenging and controversial concepts need to be introduced and explained. In essence, the idea is that for some policies, the fact that there is no tradeoff between welfare and equity will be apparent before utility-weighting. As a result, it is not necessary to apply utility weights for such policies. While utility weights may be necessary in the case of other policies, computation of the equity metrics may not be needed.[19]

First, compute the unweighted net benefits of the policy, relative to the status quo.
Second, if there are costs to the poor and the unweighted net benefit is negative or there are benefits to the poor and the unweighted net benefit is positive, <u>stop</u>. In the first instance, the status quo dominates the policy with respect to welfare and equity; and in the second instance, the policy dominates the status quo. Utility-weighting is unnecessary for determining whether or not to adopt the policy on the basis of welfare and equity because applying utility weights would only strengthen the case for or against the policy.[20]

[19] However, regardless of whether utility weights are necessary to determine whether there is a welfare/equity tradeoff, BCA will not measure aggregate welfare unless utility weights are applied.

[20] A possible exception occurs if there are modest benefits to the poor and much larger benefits to those better off. In that case, the policy would cause the gap between the two groups to widen, a decrease in equity. In this case, the

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Third, if the condition in the second step does not hold, apply utility weights, being careful to account for financial transfers.
Fourth, if the sign of net benefits changes when utility weights are applied, stop. Either the policy dominates the status quo with respect to both equity and welfare (in the first instance mentioned in step 2), or the status quo dominates the policy (in the second instance mentioned in step 2).
Fifth, if the condition in the fourth step does not hold, compute the relevant equity metrics and use them to assess the tradeoffs between welfare and equity.

4.3. An illustration of using the steps

We now present an application of the five steps to a set of real-world BCAs. These are very simple in distributional terms. We address practical complications in the following subsection. We draw three examples from a review of 26 BCAs of welfare-to-work programs (Boardman et al., 2018), which were evaluated by random assignment by MDRC, a New York-based evaluation firm. The programs provided information that was helpful for deciding whether they should continue to operate. Because these programs were all targeted at single-parent welfare recipients, distributional impacts are obviously of especial importance. The BCAs of these programs divide those affected into two groups: participants who received services under the programs, and non-participants, a category that includes everyone else in society. The key BCA components for participants were impacts on earnings, which were usually positive, and changes in government benefits, which were typically negative; the major components for non-participants were the taxes required to operate the program and reductions in the taxes needed to support the welfare system. Benefits and costs that were estimated for both participants and non-participants were divided by the number of participants, so that the numbers reported can be thought of as the net benefit generated for members of each group by a single typical participant going through the program, which allows for direct comparison of the impacts across groups.

Because the programs are small relative to the population, the incomes of non-participants should be close to the national median. For purposes of the illustration, the median incomes of participants are assumed to be at the 25th percentile, which is probably a reasonable approximation, given that participants entered the program when they were on welfare.[21] The findings from the meta-analysis describe earlier suggested setting the elasticity of the marginal utility of income, ε, at 1.61 in conducting utility-weighting, which implies a utility weight of 2.51 at the 25th percentile level and a weight of 1 at the 50th percentile level.

Our first example pertains to the GAIN program, which, beginning in 1987, operated in Riverside, California. The BCA of this program estimated that the unweighted net benefits of participants and non-participants in 2021 dollars were \$2,951 and \$5,718, respectively, per participant. Total unweighted net benefits were obviously positive at \$8,669 per participant. As indicated by the second step, utility-weighting is unnecessary because both participants

determination of whether, on balance, the equity impact outweighs the welfare impact would need to be made by direct inspection of the benefits to each group.

[21] In ongoing work, we are developing methods for more precisely determining where the different groups affected by a policy or program are located in the income distribution.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

and society as a whole received positive net benefits, implying that the policy passes the BCA test without utility weights on an *a fortiori* basis. The unweighted net benefit number is enough to conclude that the program dominates the status quo with respect to both welfare and equity. Indeed, utility-weighting is often not required in practice. For example, in 16 of the reviewed 26 BCAs, including Riverside GAIN, there were either costs to the participants and the total unweighted net benefit was negative or there were benefits to the participants and the total unweighted net benefit was positive, meaning that in all 16 cases, utility-weighting was not necessary to determine that there was no tradeoff between welfare and equity.

The second example is for a so-called jobs first program, which began in 1991, and also operated in Riverside, but less successfully than Riverside GAIN. Participants had estimated unweighted net costs of $2,238 in 2021 dollars, although because of reductions in welfare payments, non-participants had estimated unweighted net gains of $2,605 per participant. Consequently, there were small total unweighted net benefits of $368. As indicated by the third step, utility-weighting is necessary for assessing the program because there are costs to participants, yet total net benefits are positive, which is to say, on the basis of unweighted BCA, it is not possible to determine whether there is a welfare/equity tradeoff. Not surprisingly, given the small size of the estimate of total net benefits, they become negative – at $852 – once they are utility weighted. Given the sign flip, the fourth step implies that the fifth step, explicit consideration of equity, is unnecessary. With utility-weighting, the status quo is superior with respect to both welfare and equity. There is no tradeoff.

Our final example is a welfare-to-work program in which equity considerations are quite relevant: a GAIN program that began in 1988 in Alameda, California. Like Riverside GAIN, participants were found to be better off by the BCA, with unweighted net benefits estimated at $1,698 in 2021 dollars, but unlike Riverside GAIN, the BCA also found that there were unweighted net costs to non-participants of $5,788 per participant. Consequently, the unweighted loss to society was estimated to be $4,090, a reduction in potential Pareto efficiency. In this case, as per step 3, and as in the second example above, utility-weighting is necessary in order to determine whether there is a welfare/equity tradeoff. After utility-weighting, net benefits for participants increase to $2,734, net costs for non-participants remain at $5,788 per participant, and total net costs shrink to $3,055. Because the sign of net benefit does not change, we can conclude that there is a welfare/equity tradeoff.

In order to assess the acceptability of the tradeoff, it is necessary to move on to the fifth step and compute the equity metrics we propose in section 4. The distributional impact (the reduction in the welfare gap between the two groups) is the absolute sum of the utility-weighted benefit to participants and the utility-weighted cost to non-participants, which is $8,522. The welfare loss per dollar of reduction in welfare gap (what could be called the price of inequality reduction) is $.36. The poverty impact (the increase in the welfare of participants) is $2,734, and the welfare loss per dollar of poverty impact (what could be called the price of poverty reduction) is $1.1. The breakeven equity weight is 2.12 ($5,788/$2,734). Although the welfare cost per dollar of distributional impact is small, a decision maker might worry that many non-participants have incomes that are not exceptionally high. The welfare cost per dollar of poverty impact is rather high, but might still be considered tolerable, particularly given the fairly low breakeven weight, which implies that in order to prefer the program to the status quo, the decision maker would need to feel a bit over twice the concern for participants than for non-participants in order for the distributional benefits of the program to outweigh the welfare loss. This example demonstrates what we mean when

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

we say that our metrics allow for *quantitative* assessment of welfare/equity tradeoffs. They put the comparison of welfare impact and equity impacts into intuitively meaningful quantitative terms with which decision makers can draw their own conclusions about the pros and cons of the programs.

As part of a multigoal analysis, the decision maker would have other concerns besides the distribution of costs and benefits among income groups. For example, in the Alameda GAIN program, operating costs were estimated to be $10,893 per participant in 2021 dollars, which is large for welfare-to-work programs. The decision maker would presumably multiply that figure by the expected number of future participants to determine if there is sufficient space in the government's budget. As this discussion suggests, while utility-weighting is helpful, unweighted estimates of disaggregated benefit and costs are also needed for decision-making.

4.4. Practical concerns with the application of utility-weights

The welfare-to-work BCAs described in the previous subsection presented rather non-challenging illustrations of how the steps listed in section 4.2 could be implemented. Many BCAs will be more challenging. In this subsection, we briefly mention some of these challenges. We do not attempt to detail how to meet these challenges, as doing so is beyond the scope of this paper and is a topic of our ongoing research.

One advantage that the welfare-to-work BCAs provided is that separate BCAs were conducted for the low-income program participants and the relatively higher income non-participants, a framework that was developed by Kemper, Long, and Thornton (1981) and has subsequently been used in most BCAs of government programs that invest in the low-income population. As a result, it was immediately clear whether the studied programs made one group better off and the other group worse off, and consequently whether utility-weighting was necessary. While many other BCAs involve programs and policies that affect both lower and higher income households (e.g., those of health, safety, and environmental regulations), information is often not readily available on where in the income distribution those affected fall. Hammitt, Robinson and Zeckhauser (2016), for example, point out that while the data often exist to determine how the benefits of health, safety, and environmental regulations are distributed by income, it is extremely difficult to determine the distribution of costs because the affected firms may raise prices, reduce wages, or simply accept reduced profits, and such changes may affect different groups differently. Once distribution is determined, two or more income groups must be defined. Then, the estimated benefits and costs must be allocated among these groups. It may be especially difficult to determine how consumer surplus is distributed among groups when it is estimated through aggregate demand curves. Additionally, whereas financial transfers – including tax-and-transfer programs, transfers from employers to employees in the form of wages, and transfers from consumers to firms in the form of prices – can be ignored in unweighted BCA, they must be identified and accounted for before applying utility weights. Doing all this is an obvious precursor to implementing steps 2 through 5. We recognize that these steps involve non-trivial practical challenges. We address a few here, but have chosen to leave a more fulsome consideration of practicalities to future research.

The measure of income that is used in distributional analysis should ideally be the income available for consumption. Consequently, it should be net of taxes and should include government transfers, non-monetary benefits, and possibly even borrowing and saving, and

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

should possibly take account of the fact that the cost of living varies across geographic regions. In practice, estimating the ideal measure of income may prove infeasible, and an approximation may be necessary. The question of whether to use per-capita or household income also arises. One approach is to make "equivalizing" assumptions about the distribution of consumption across household members, taking account of things like economies of scale in household production. An example of how this can be done in practice can be found in the U.K. Greenbook (Her Majesty's Treasury, 2018).

Another important consideration that will need to be addressed in some BCAs is how to handle valuations that are, in practice, typically based on averages across income levels, such as the value of statistical life, quality-adjusted life years, the value of time, and those estimated with contingent valuation techniques. There are two possible approaches to treating benefits that are averaged across income groups. The first is use the averages and not weigh them. Banzhaf (2011) has argued, for instance, that using averages across income levels represents an approximation to utility-weighting because it implicitly assumes that individuals with different incomes would express the same valuation if they had the same income. The second approach is to first assign individuals the amount they would be willing to pay for the benefit, which is presumed to usually rise as income increases, and then utility weight the resulting values.[22] Our own preliminary efforts concerning the value of statistical life imply that these two approaches produce very different results. For example, there is a considerable literature on the income elasticity of the value of a statistical life. (Hammit and Robinson, 2011) If the elasticity of VSL is significantly different from the elasticity of marginal utility of income, using average VSL will generate significantly different results than first computing VSL at different income levels using the income elasticity of the VSL and then applying utility weights using the elasticity of marginal utility of income. Contingent valuation studies that account for variation by income provide another example with the same practical considerations.

Finally, although our welfare/equity ratios do not strictly require that society be categorized into only two groups (for example, the denominator can the Gini or Atkinson coefficients), the ratio of welfare to what we are calling "welfare gap," as well as the breakeven equity-weight, do. In many cases, such as the kinds of programs in our worked example above, the bifurcation will be relatively obvious. In others it may be necessary to use a relatively crude bifurcation, perhaps something like above and below median income. We believe that this kind of bifurcation, while not ideal, will nonetheless be informative in many cases. And, as with many other aspects our recommendations, the effect on the metrics of the definition of groups can be explored in sensitivity analysis.

Partially because of the issues just discussed and partially because of lack of agreement over the appropriate weights, both utility and equity weighting are rare in actual benefit-cost analyses (Adler, 2013). One apparent exception is in global studies of climate change where the disparity of incomes among countries is enormous (Nurmi and Ahtiainen, 2018). In addition, distributional weighting has occasionally been conducted in BCAs of programs specifically targeted at the low-income population (for example, Greenberg, 2013 and Gubits et al., 2018). Even determining how the benefits and costs of regulation policies are distributed among different income groups appears infrequent. For example, Robinson,

[22] Approaches for doing this for the value of statistical life and the value of time can be found in Boardman et al., 2018, Chapter 17.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Hammitt, and Zeckhauser (2016) examined 24 policies that estimated the benefits and costs of federal health regulations and "found that these analyses pay relatively little attention to distribution" (p. 323). They also mention that several recent reviews do not find studies that determine the distribution of "both the costs and benefits of more conventional energy and environmental regulations" (p. 309). Without such information, weighting is obviously infeasible without considerable assumptions.

Section 5: Conclusion

Our thesis is two-fold: First, using utility weights to correct for diminishing marginal utility of income is a technical solution to what is essentially a methodological problem: WTP is the only viable metric for BCA, but does not correctly measure welfare because it is biased against the poor. In the context of multi-goal analysis, welfare is one of the criteria about which decision makers need information, and it must be measured in an unbiased way. Thus, utility weights must be applied, to correct for diminishing marginal utility of income. Utility-weighted BCA has an intuitively comprehensible interpretation: it is the aggregate WTP for the costs and benefits of a policy that we would observe if all members of society had income at the current median. Because the decision of what utility weights to use is a technical decision, rather than a moral judgment, it belongs under the purview of analysts. There seems to be a strong evidentiary basis for a consensus on the correct weights.

Our second thesis is that the use of equity weights in BCA, in an attempt to account for what has been referred to as "diminishing moral value of welfare" (Adler 2008), undermines the ability of decision makers to assess tradeoffs between welfare and equity. We begin with the premise that the appropriate way to make decisions that involve different decision-making criteria, or goals, is multi-goal analysis, in which is it necessary to provide separate information about each criterion, so that decision makers can apply their own values to the assessment of tradeoffs among criteria. This premise leads directly to the conclusion that equity-weighting must be avoided, as it conflates welfare and equity and thus contains no information about welfare or equity separately, as required in multi-goal analysis. The same equity-weighted net benefit can be generated by a policy that increases aggregate welfare while decreasing equitability and another policy that decreases aggregate welfare while increasing equitability. There is no way for decision makers to assess the tradeoffs between the two. Nonetheless, intuitively useful information about the tradeoffs between welfare and equity can be extracted from a utility-weighted BCA by calculating the equity metrics we have proposed in section 3. Provided that some quantitative analysis of equity impacts has been included in multi-goal analysis, these metrics allow decision makers to make intuitively comprehensible, quantitative assessments of the tradeoffs between welfare and equity created by a policy.

We derive these conclusions about utility-weighting and equity-weighting by the application of a simple set of logical steps to the initial premise that regulatory review, and indeed all analysis of policies that affect multiple social goals, should be conducted using multi-goal analysis. We acknowledge that this premise is contestable. We cannot "prove" that it is "correct." However, we believe our defense of the premise will prove compelling to many in the field of benefit cost analysis, as it has been to the authors of the primary textbook on the subject (Boardman et al., 2018).

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

We distill our analysis into a set of recommendations which we present and explicate in section 4, along with a set of steps for assessing welfare/equity tradeoffs with the least information possible. We offer these recommendations as a direct response to the Biden memorandum on "Modernizing Regulatory Review," and believe that the practical challenges posed by the recommendations can be overcome well enough to make regulatory review more informative, in ways that can promote equity. And we believe the same is true with respect to any context in which BCA is used as an aid to decision making.

Much of what we present here is based on established findings that are not novel. However, we believe that we have made a number of novel contributions. First, we establish that utility-weighting is a technical matter in the measurement of welfare, and thus should be addressed by analysts. We recognize that in practice, analytical choices of this sort are often imposed from above in ways that are politically influenced, but this does not alter the fact that utility-weights do not involve value judgements, and that the choice of weights ought properly to be established by a consensus of economists and analysts, as Marshall implicitly recommended in 1885. Second, we have pointed out that, contrary to a widely held view, there is a compelling basis for such a consensus, based on the model of isoelastic income utility – which is universally adopted in the literature on diminishing marginal utility of income – and a thorough meta-analysis of the very large number of estimates of its key parameter. Third, we have refuted the defense of unweighted BCA based on the idea that it will benefit both wealthy and poor in the long-run by pointing out that it will, nonetheless benefit the wealthy more than the poor in terms of welfare, and thus increase welfare inequality, even if it makes the poor better off over time. Fourth, we have observed that equity-weighting violates the requirements of multi-goal analysis and thus makes it impossible for decision makers to apply their own value systems to the assessment of tradeoffs between welfare and equity. Fifth, though breakeven distributional weights are not new, we are not aware of any application of what we are calling welfare/equity ratios, or the idea that the optimal way to assess the acceptability of tradeoffs between welfare and equity in regulatory review is cost-effectiveness analysis. Finally, we have offered a simplified set of steps for determining whether a policy actually creates a tradeoff between welfare and distributional impact, and for assessing such tradeoffs when they exist, using the least information possible.

Acknowledgements. We gratefully acknowledge the extremely helpful comments of Dave Weimer, Clark Nardinelli, Samantha Yi, and the participants of the 2022 conference of the Society for Benefit Cost Analysis, as well as the comments of anonymous reviewers.

Supplementary material. The supplementary material for this article can be found at http://doi.org/10.1017/bca.2023.5.

Competing interest. The authors declare none.

References

Acland, Daniel and David Greenberg. 2023. *Weights for Distributional Weighting and Values for Social Discounting: A Meta-Analysis of the Elasticity of Marginal Utility of Income Estimates*. Unpublished manuscript. University of Maryland, Baltimore County and University of California, Berkeley.

Adler, Matthew D. 2008. "Risk Equity: A New Proposal," *Harvard Law Review 32*:1–47.

Adler, Matthew D. 2013. "Cost-Benefit Analysis and Distributional Weights: An Overview," *Duke Environmental and Energy Economics Working Paper EE*, 13–04.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Adler, Matthew D. 2019. *Measuring social welfare: An introduction*. Oxford University Press.

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

Banzhaf, H. Spencer. 2011. "Regulatory Impact Analysis of Environmental Justice Effects," *Journal of Land Use and Environmental Law 27(*1).

Bardach, Eugene and Eric M. Patashnik. 2019. *A Practical Guide for Policy Analysis: The Eightfold Path to More Effective Problem Solving* Washington, D.C.: CQ Press.

Belton, Valerie and Theodor Stewart. 2002. *Multiple Criteria Decision Analysis: An Integrated Approach*. Springer Science & Business Media.

Boardman, Anthony E., David H. Greenberg, Aidan R. Vining, and David Weimer. 2018. *Cost-Benefit Analysis: Concepts and Practice*. 5th ed. New York, NY: Cambridge University Press.

Brent, Robert. J. 2017. *Advanced Introduction to Cost–Benefit Analysis*. Cheltenham, UK: Edward Elgar Publishing.

Cowell, Frank A. and Karen Gardiner. 1999. *Utility Weights* STICERD. London, UK: London School of Economics. http://darp.lse.ac.uk/papersDB/Cowell-Gardiner_(OFT).pdf.

Dasgupta, Partha, Amartya Sen, and S. Marglin. 1972. *Guidelines for Project Evaluation*. New York: UNIDO, United Nations.

Eckstein, Otto. 1961. A Survey of the Theory of Public Expenditure Criteria. In *Public Finances: Needs, Sources, and Utilization:* 439–504. Princeton: Princeton University Press.

Farrow, Scott. 2011. "Incorporating Equity in Regulatory and Benefit-Cost Analysis Using Risk-Based Preferences," *Risk Analysis: An International Journal*, *31*(6): 902–907.

Farrow, Scott. 2021. *On Balance: When All Lives Matter Equally: Equity Weights for BCA by Combining the Economics of VSL and US Policy*. Society for Benefit-Cost Analysis Blog.

Feldstein, Martin. 1974 "Distributional Preferences in Public Expenditure Analysis" In Hochman, H. M. and G. E. Peterson (Eds) *Redistribution Through Public Choices*. New York: Columbia University Press, 136–161.

Fleurbaey, Marc. 2010. "Assessing Risky Social Situations," *Journal of Political Economy 118*(4): 649–680.

Frey, Brono. S., Simon Luechinger, and Alois Stutzer. 2010. "The Life Satisfaction Approach to Environmental Valuation," *Annual Review of Resource Economics* Vol. *2*:139–160.

Graham, John D. 2008. "Saving Lives Through Administrative Law and Economics," *University of Pennsylvania Law Review: 157*(2):395–540.

Greenberg, David H. 2013. A Cost-Benefit Analysis of the Tulsa IDA Program. *Journal of Benefit-Cost Analysis*, *4* (3): 263–300.

Groom, Ben and Maddison, David. 2019. "New Estimates of the Elasticity of Marginal Utility for the UK," *Environmental and Resource Economics*. http://doi.org/10.1007/s10640-018-0242-z.

Gubits, Daniel et al. 2018. *BOND Implementation and Evaluation: Final Evaluation Report, Volume 1*, Washington, DC: Social Security Administration

H.M. Treasury 2018. *Green Book: Appraisal and Evaluation in Central Government*, London: HMSO.

Hammitt, James K. and Lisa A. Robinson. 2011. "The Income Elasticity of the Value per Statistical Life: Transferring Estimates between High and Low Income Populations," *Journal of Benefit-Cost Analysis 2*(1).

Harberger, Arnold C. 1978. "On the Use of Distributional Weights in Social Cost-Benefit Analysis," *Journal of Political Economy 86*(2): S87–S120.

Havranek, T., Horvath, R., Irsova, Z. and Rusnak, M. 2013. *Cross-Country Heterogeneity in Intertemporal Substitution*. William Davidson Institute Working Paper Number 1056. Ann Arbor Michigan: University of Michigan.

Kagan, S. 2018. *Normative Ethics*. New York: Routledge.

Kemper, Peter, David A. Long, and Craig Thornton. 1981, *The Supported Work Evaluation Final Report*. New York: MDRC.

Krutilla, Kerry 2005. "Using the Kaldor-Hicks Tableau Format For Cost-Benefit Analysis And Policy Evaluation," *Journal of Policy Analysis and Management: The Journal of the Association for Public Policy Analysis and Management*, *24*(4), 864–875.

Landemore, Helene. 2017. "Deliberative Democracy as Open, Not (Just) Representative Democracy," *Dædalus, the Journal of the American Academy of Arts & Sciences.*

Layard, Richard, Guy Mayraz, and Stephen Nickell. 2008. "The Marginal Utility of Income," *Journal of Public Economics 92*:1846–1857.

Lefant, Jean-Sébastien. 2012. Indifference Curves and the Ordinalist Revolution., *History of Political Economy*, *44*(1):113–155.

Little, Ian M. D. and James A. Mirrlees. 1974, *Project Appraisal and Planning for Developing Countries* London, U.K.: Heinemann Educational.

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press

Marshall, A. (1885). *The present position of economics, an inaugural lecture*. Macmillan.
Mishan, E. J., 1988. *Cost-Benefit Analysis*. London, UK: Unwin Hyman Ltd.
Nurmi, Vaino and Heini Ahtiainen. 2018. "Distributional Weights in Environmental Valuation and Cost-benefit Analysis: Theory and Practice," *Ecological Economics 150*:217–228.
Office of Management and Budget (2003) "Circular A4, Regulatory Analysis." Federal Register, 68 FR 58366
Ray, Anandarup. 1984. *Cost-Benefit Analysis: Issues and Methodologies*. Baltimore: John Hopkins University Press.
Robinson, Lisa A., James K. Hammitt, and Richard J. Zeckhauser. 2016. "Attention to Distribution in US Regulatory Analyses," *Review of Environmental Economics and Policy.*
Scarborough, Helen and Jeff Bennett. 2012. *Cost-Benefit Analysis and Distributional Preferences: A Choice Modelling Approach* Cheltenham, UK: Edward Elgar Publishing.
Sheriff, Glenn and Kelly B. Maguire. 2020. "Health Risk, Inequality Indexes, and Environmental Justice," *Risk Analysis*, *40*(12), 2661–2674.
Sunstein, Cass R. 2018. *The Cost-Benefit Revolution* Boston: MIT Press.
U.S. Census Bureau. 2021. American Community Survey Public Use Microdata Samples.
Vining, Aidan R. and Boardman, Anthony E. 2006. "Metachoice in policy analysis," *Journal of Comparative Policy Analysis*, *8*(01), pp.77–87.
Weimer, David L., and Aidan R. Vining. 2017. *Policy Analysis: Concepts and Practice*. New York: Routledge.
Weisbrod, Burton A. 1968. "Income Redistribution Effects and Benefit-Cost Analysis," In S. B. Chase (Ed.) *Problems in Public Expenditure Analysis:* (177–208) Washington, D.C.: The Brookings Institution.
Wiener, J. B., Bronsteen, J., Buccafusco, C., & Masur, J. S. (2013). Well-Being Analysis vs. Cost-Benefit Analysis.
Yitzhaki, Shlomo. 2003. "Cost-Benefit Analysis and the Distributional Consequences of Government Projects," *National Tax Journal 56*(2):319–336.

Cite this article: Acland, D. J. and Greenberg, D. H. 2023. "Distributional weighting and welfare/equity tradeoffs: a new approach." *Journal of Benefit-Cost Analysis* 14: 68–92, doi:10.1017/bca.2023.5

https://doi.org/10.1017/bca.2023.5 Published online by Cambridge University Press