CAMBRIDGE
UNIVERSITY PRESS

ARTICLE

Which values should be built into economic measures?

S. Andrew Schroeder

Department of Philosophy, Claremont McKenna College, 850 Columbia Ave, Claremont, CA 91711, USA
Email: andrew.schroeder@cmc.edu

Abstract
Many economic measures are structured to reflect ethical values. I describe three attitudes
towards this: *maximalism*, according to which we should aim to build all relevant values
into measures; *minimalism*, according to which we should aim to keep values out of
measures; and an *intermediate* view. I argue the intermediate view is likely correct, but
existing versions are inadequate. In particular, economists have strong reason to structure
measures to reflect *fixed*, as opposed to *user-assessable*, values. This implies that, despite
disagreement about precisely how to do so, economists should standardly adjust QALYs
and DALYs to reflect egalitarian values.

1. Values in economic measures: three positions

It is well-known that many economic measures are structured to reflect social,
ethical, or political values, or have such values 'built in' to them. Take, for example,
the unemployment rate. As standardly calculated in the USA a number of groups
are excluded from the calculation, counting as neither employed nor unemployed:
those younger than 16, those not actively looking for work, those in prison, and
residents of certain health-care facilities. Many of these individuals have a clear
employment status, so there is no technical reason for their exclusion. It seems more
likely that in at least some of these cases the exclusion is based on the ethical
judgement (whether correct or not) that members of these groups in some sense
don't deserve jobs, or have no claim on society to provide them with jobs. I suspect,
for example, that many people think that those younger than 16 generally ought
not work and therefore can have no claim on society for a job, that incarcerated
individuals have forfeited any claim to a job for the duration of their incarceration,
and that it is so challenging to find a job for full-time residents of health-care
facilities that society has no obligation to provide employment.[1] Similar claims could

---

[1]To be clear, I do not mean to be endorsing these claims. I strongly suspect, however, that views like these
lie behind at least some exclusions from unemployment calculations.

be made about how many other economic measures, including GDP, price indices and health economic measures such as quality-adjusted and disability-adjusted life years (QALYs, DALYs), are structured to reflect ethical or political values.[2]

What should we think about adjusting or structuring measures for such ethical reasons? There has been a fair amount of discussion on narrow questions – questions about specific value choices that structure or could structure measures. There has been debate, for example, about whether incarcerated individuals really should be excluded from unemployment calculations, about whether QALYs should give less weight to infant deaths, and about whether luxury goods should be excluded from price indices. Oftentimes, this debate proceeds by directly evaluating the ethical issue at stake. (Have prisoners really forfeited any claim to a job? Are infant deaths morally less significant than adolescent deaths?) But it seems to me that before engaging with these narrow questions, we must first address a set of broader questions which may render the narrower ones irrelevant: should the unemployment rate even be adjusted for differential claims to jobs in the first place? Should a putative measure of health like the QALY even try to reflect the moral significance of different types of health losses? These broader questions have received less explicit attention.

One way to reject these broader questions would be to argue that they don't really make sense, since *any* choice we make will reflect some ethical value or other. Just as giving less weight to infant deaths in QALY calculations would reflect the view that infant lives are less important than adolescent lives, assigning them equal weight would reflect the view that they are of equal importance. Elsewhere, I explain why I reject this view: although there may be a sense in which all measures are thoroughly infused with values, there is a distinct and important sense in which choosing to discount infant lives for ethical reasons involves values in a way that simply weighing them equally need not (Schroeder 2017a). In other words, it is possible for a health economist to calculate mortality statistics while remaining agnostic on the relative moral importance of infant vs. adolescent lives. In this essay, I will take for granted that this is correct – that there is an important sense in which some choices, including those mentioned earlier, render an economic measure value-laden in ways it otherwise needn't be. I think this is the standard view amongst economists, and in any case it is presupposed by the authors I will go on to discuss.

Given, then, that economists at least sometimes have a choice about whether to adjust measures to reflect ethical values, we can identify three broad stances one could take towards doing so, each of which has some support among economists and philosophers of economics. First, we can imagine a *maximalist* position, according to which, ideally, every value relevant to the associated policy decision(s) should be built in to an economic measure. If, for example, policy-makers are called upon to minimize unemployment *among those with claims to jobs*, then they should be provided with unemployment data that excludes those without claims to jobs. And if health ministers should give more weight to adolescent deaths than to infant deaths, QALY calculations should be weighted accordingly. Alan Williams offers a defence of a view like this:

---

[2]On these points see Stiglitz *et al.* (2010), Voigt (2012), Reiss (2013: Ch. 8; 2017: §3.2) and Schroeder (2017a).

If the nature and implications of particular positions are to be clarified in a policy-relevant way, this discussion has to move on to seek quantification of what are otherwise merely vaguely appealing but ambiguous [ethical] slogans ... Only with some quantification will it be possible to devise rules that can be applied in a consistent manner with a reasonable chance of checking on performance. (Williams 1997: 120; cf. Menzel *et al.* 1999; Nord 1999; Schokkaert 2015).

Though a number of economists appear to endorse maximalism or hold views substantially similar to it, it is not very popular.

Much more popular has been what I will call *minimalism*. In some cases, it may not be possible to avoid value judgements when structuring a measure. Calculating any sort of unemployment rate, for example, requires determining what counts as employment, and this will involve value judgements. (Is someone who works one hour a week employed? What about an intern who receives a parking subsidy but no other material compensation?) But, the minimalist says, we should structure our measures to reflect value judgements only when necessary; whenever possible, we should keep ethical values out of economic measures. Economic measures, after all, are supposed to be scientific, and scientific measures are generally expected to be value-free.[3] It would seem incredibly odd to adjust other scientific measures to reflect ethical considerations – to exclude or discount voluntary, informed smokers from lung cancer statistics, for example, on the grounds that their illness should be a lower public health priority. Similarly, perhaps we should let the unemployment rate simply measure unemployment, and QALYs simply measure health. Judgements about whose unemployment matters or about whose health problems are of greater social concern should be left to policy-makers, not made by economists. Christopher Murray and colleagues seem to endorse a minimalist view when they suggest that DALYs should be 'viewed as a strict summary measure of population health', and therefore not adjusted to account for certain ethical factors (Murray *et al.* 2012: 16). And the US Bureau of Labor Statistics insists that its data 'must satisfy a number of criteria, including ... impartiality in both subject matter and presentation'.[4]

Having laid out the two extreme positions, we are left with a collection of *intermediate views*, which hold that some but not all policy-relevant values should be built in to economic measures. Any proponent of an intermediate view, however, faces an obvious question: if some but not all policy-relevant values are to be built in to a given measure, how should economists decide which values to build in and which to leave out? Though in practice intermediate views seem to be fairly common, economists rarely give an explicit answer to this question. Murray, however, offers a proposal which I think is consonant with many economists' practices:

If many individuals after deliberation hold a preference or value then this value should be considered seriously. We should investigate ... the likely reasons why many individuals hold such a view. If these reasons appear to be persuasive and

---

[3]I think we have clear evidence that scientific measures, and scientific practices more generally, are not and should not be value-free (Douglas 2009; Reiss 2017). But that remains the (mistaken) perception.

[4]See FAQ #1 at https://www.bls.gov/bls/faqs.htm.

do not contravene important "ideal-regarding principles," these preferences should be incorporated into the construction of DALYs. (Murray 1996: 5)[5]

Murray seems to be proposing two criteria. A policy-relevant value should be incorporated into the DALY if and only if that value (1) is widely held among those who have reflected on it, and (2) can be supported by good reasoning (or at least does not conflict with important moral principles). We could, somewhat loosely, sum this up by saying that Murray wants to build uncontroversial values, and only uncontroversial values, into the DALY – and, presumably, other economic measures.

This, I think, seems like a sensible view, which tries to accommodate the insights lying behind both maximalism and minimalism. In permitting some values to be built in, this view accepts the maximalist's insight that quantifying policy-relevant values and building them into our measures will enable policy-makers to more effectively pursue their goals. But, in restricting those values to uncontroversial ones, the economist doesn't seem to be stepping beyond her expertise in any significant way, or usurping the role of the policy-maker. After all, the only values being built in are ones that are the subject of general agreement and which can be supported by good reasoning.

Nevertheless, I think Murray's view is mistaken. In the succeeding sections, I'll argue by way of a case study that Murray's proposal fails because it looks only at properties of the value choice itself – how widely held it is, and what substantive reasons support it – while failing to look at the way in which the value is built in to the measure. Some value choices are built into measures in what I will call *fixed* ways, while others are *user-assessable*. I will argue that economists generally have stronger reason to build fixed value choices into a measure, even if they are more controversial than user-assessable value choices. In the end, I will suggest that while Murray was probably right to endorse an intermediate view over maximalism or minimalism, an adequate version of the intermediate view will be much more complicated than Murray's proposal. I will conclude by briefly weighing in on the ongoing debate about whether measures of health should standardly be adjusted to account for inequality, arguing that they should.

2. Case study: values in the DALY

DALYs, a relative of QALYs, are a composite measure of morbidity and mortality designed to measure the amount of health lost to a particular cause or event. Their most prominent use is as the main outcome measure in the influential Global Burden of Disease Study, of which Christopher Murray is one of the lead authors. Though the calculation of DALYs is complicated, for our purposes we need only look at the mortality component, which is fairly straightforward. The health loss attributed to a premature death is calculated based on the years of life lost to that death. Suppose, in a country where life expectancy at birth is stable at 80, a car accident kills a 30-year-old and a 40-year-old. That car accident took roughly 50 and 40 years of life from its victims, respectively, and so 50+40=90 DALYs would be attributed to the accident. Historically, however, many epidemiologists and

---

[5]Obviously, this quote offers a different view from the 2012 passage I cited earlier. Murray's views seem to have evolved in a minimalist direction between 1996 and 2012.

economists using DALYs have not been content to rest there. The DALYs resulting from some events have typically been adjusted to reflect various ethical values (Polinder *et al.* 2012; Schroeder 2017a). In the remainder of this section, I will discuss three of those value choices, to see what they can tell us about the intermediate view Murray endorses.

The first value choice is that involved in *age-weighting*. Most people believe that each year of life is not of equal value from a political or ethical perspective. A surgery that would extend the life of a young adult by two years should be a higher policy priority than one that extends the life of an infant by two years. And most people, if forced to choose, prefer to save the life of a teenager before saving the life of an infant, despite the fact that the infant likely has more years left to live. Murray, in an article explaining the methodology behind the DALY, accepts a version of this argument:

> The well-being of some age groups, we argue, is instrumental in making society flourish; therefore collectively we may be more concerned with improving health status for individuals in these age groups. (Murray and Acharya 1997: 719)

On the basis of empirical investigations of age-preferences among the public, the original DALY assigned extra weight to health losses between the ages of 9 and 54, and reduced weight to health losses outside of that range. Many have criticized this value judgement, some arguing that different age-weights should be considered (Barendregt *et al.* 1996; Tsuchiya 2000), others arguing that no age-weights should be used (Anand and Hanson 1998; Williams 1999; Bognar 2008). Murray and colleagues eventually accepted these concerns, and in recent updates to the Global Burden of Disease Study have abandoned the use of unequal age-weights (Murray *et al.* 2012).

Second, consider the standard economic practice of *discounting*. In economic calculations, the value of future costs and benefits is typically discounted by a small percentage per year. Most DALY calculations treat health in the same way, discounting the value of future health. So, a death averted this year counts for more than an otherwise-similar death averted next year. Although a variety of arguments have been offered for discounting, Murray chose an explicitly ethical one in the construction of DALYs:

> In the construction of DALYs, I have struggled with two options: to use a low positive discount rate to capture the uncertainty that increases with time *and, more importantly, to reduce the problems of excessive sacrifice* ... As with previous versions of this study, the baseline DALY measure incorporates a three per cent discount rate. (Murray 1996: 53–54; emphasis added)

As with age-weighting, the practice of discounting has been incredibly controversial, and in the most recent iteration of the Global Burden of Disease Study, Murray has abandoned it (Murray *et al.* 2012) – though DALYs are discounted in many other studies (Polinder *et al.* 2012).

A third ethical value built in to the DALY is connected to the way it handles *life expectancy*. To calculate the years of life lost to a premature death, we need to assign a life expectancy to the victim. The most obvious and natural way to do that is to ask, of the victim, how many more years she would have lived, had she not died of the cause in question (Williams 1999: 4). Call this a *counterfactual* approach to life

expectancy (CLE). Murray, however, points out a problem with CLE: suppose two 40-year-olds, one in Kigali and the other in Paris, die in car accidents.

> [Using CLE] would lead us to conclude that the death of a 40 year-old woman in Kigali contributes less to the global burden of disease than the death of a 40 year-old woman in Paris because the expectation of life at age 40 is lower in Rwanda than in France. Equivalent health outcomes would be a greater burden in richer communities than in poorer communities. (Murray 1996: 14)

This seems unfair. Certain populations are already disadvantaged by low life expectancy, and then deaths in those populations register fewer DALYs because of that disadvantage. When measuring the burden of disease on a population, it seems perverse to count a death for less because the victim would otherwise have died of another health problem shortly thereafter.

What, then, should we do? Murray proposes that we apply the same life expectancy to all deaths at a given age anywhere in the world, making the death of every 40-year-old woman count the same, regardless of her personal characteristics. This is, clearly, a value choice: this proposal was adopted because the apparently value-free route – to simply count how many years were actually lost due to the death – yielded (by Murray's lights) morally unacceptable results. Murray applies the same life expectancy to all deaths not because all people *do* live to roughly the same age, but because he believes that they *should*. Put another way, the DALY does not measure how many more years a victim *would have* lived; it measures how many more years she *should have* lived. Call this an *aspirational* approach to life expectancy (ALE). Murray explicitly says that he chose to use ALE because of its 'egalitarian nature' (1996: 15) – a dimension we can highlight by noting that calculating DALYs using ALE is equivalent to calculating them using CLE, while giving extra weight to years of life lost by those disadvantaged with low life expectancy (Williams 1999: 5).

Unlike age-weighting and discounting, the use of aspirational life expectancy has been almost entirely uncontroversial. As far as I can tell, since Murray's original discussion (Murray 1996) there has been only one short book chapter dedicated to discussing the life expectancy chosen for the DALY (Vos 2002) – in contrast to the large number of articles written about age-weighting and the nearly limitless number on discounting. A few articles do note the use of aspirational life expectancy, but most either don't recognize it as a value choice or else endorse the value choice.[6] Not surprisingly, current versions of the Global Burden of Disease Study continue to incorporate a version of aspirational life expectancy.

If age-weighting and discounting have been so controversial, why has Murray's choice to use an ethically loaded conception of life expectancy largely escaped criticism? There are several possible explanations. Perhaps the substantive ethical arguments in favour of aspirational life expectancy are stronger than the arguments for age-weighting and discounting. Perhaps those ethical beliefs are more widely

---

[6]The only clear critic of aspirational life expectancy that I know of is Williams (1999). Articles which mention but don't seriously criticize the value choice include (Anand and Hanson 1997; Lyttkens 2003; Arnesen and Kapiriri 2004; Voigt 2012). Those articles focus on (and sometimes criticize) the use of different life expectancies for each gender, but otherwise accept the use of aspirational life expectancy.

accepted among philosophers, policy-makers and economists. Or perhaps many authors simply haven't noticed the value choice involved in the use of aspirational life expectancy.[7] These explanations may be part of the truth. But I think another factor may also be at work – one that would not merely explain the differential treatment, but would also justify it.

3. What's special about life expectancy?

Suppose you are a policy-maker and are given the un-age-weighted DALYs attributable to two diseases, but you think that health losses at different ages should be of different social concern. What information would you need to produce an estimate of age-weighted DALYs? You would need (1) a breakdown of the ages at which DALYs were lost, and (2) your preferred age-weighting function. On most age-weighting proposals, (2) is relatively simple, and the information needed for (1) can often be readily estimated by policy-makers, because the age profiles of many diseases are relatively constant and widely known. (Malaria primarily kills the very young; strokes mostly occur in the elderly; US car accident fatalities strike all ages, with relative spikes among young adults and the elderly; etc.) And, even if (1) isn't previously known to policy-makers, that breakdown is often published in a study like the Global Burden of Disease Study. What all this means is that, given un-age-weighted results, a moderately sophisticated policy-maker can usually produce a reasonable estimate of what age-weighted results would look like, without needing access to the raw data. Much the same, I think, can be said about discounting. Call value choices like this, which can be adjusted for after-the-fact by decision-makers, *user-assessable*.

Now, compare that to the case of life expectancy. Suppose you are given DALYs calculated using counterfactual life expectancy, but, moved by Murray's arguments, you would prefer to base your policy decisions on DALYs calculated using aspirational life expectancy. What information do you need? Life expectancy varies significantly by age, gender and geography (among other factors), so at minimum you need DALYs broken down along all of those dimensions.[8] Then you need data on actual life expectancy for each age-gender-geography triple, so that you can determine the difference between the CLE used in the results you were given and the ALE you prefer. These differences are not always as predictable as one might suspect. (For example, although richer countries typically have higher life

---

[7]There is probably something to this. Although in many works (e.g. Murray 1996) the architects of the DALY were clear about the value choice involved in life expectancy, they tended to place much greater emphasis on age-weighting and discounting, going so far as to build 'modulation factors' into the DALY equation, allowing them to calculate DALYs without age-weighting and with different discount rates. Accordingly, they officially refer to 'DALY [*r, K*]', where *r* indicates the discount rate used and *K* is binary variable indicating the presence or absence of unequal age-weights. All of this, I think, has led many to erroneously conclude that DALYs incorporate only three value choices – disability weights (not discussed in this article), age-weights and a discount rate. (See e.g. Anand and Hanson 1998.)

[8]Depending on how fine-grained life expectancy is to be assessed, we might also want information on race, economic status, insurance status, co-morbidities, etc. I will set aside such considerations in the main text, though, since I think the question of how narrowly to assess life expectancy raises additional issues, beyond those considered in this paper.

expectancy than poorer countries, there are plenty of outliers.) Adjusting life expectancy from CLE to ALE (or vice versa) therefore requires much more data – essentially, the unaggregated data – as well as a much more complex calculation. The value judgements around life expectancy therefore seem much harder for a user to implement on her own, compared with those connected to age-weighting or discounting. Call value choices which can't realistically be applied by a user to a summary measure, e.g. due to high informational requirements or computational complexity, *fixed* value choices.

As further evidence that the values in the DALY connected to age and time are much more easily user-assessable than those connected to life expectancy, consider a well-known 2004 article in *Health Policy* by Trude Arnesen and Lydia Kapiriri. Arnesen and Kapiriri set out to determine to what extent the final results in the Global Burden of Disease Study were sensitive to the value judgements built in to the DALY. Concerning age-weighting and discounting, they ultimately showed that those value choices 'have a decisive impact on the relative distribution of the burden between age groups' (Arnesen and Kapiriri 2004: 145). However, they add, 'We were not able to assess the effect of alternative life expectancies. We would then have needed access to the details in the epidemiological information, as well as life tables with hypothetical life expectancies in addition to life tables based on observations' (2004: 146). Two trained specialists, then, attempting to assess the impact of value choices on a major study found it relatively easy to recalculate DALYs with alternative age-weighting and discounting functions, but were unable to do the same for life expectancy.[9] That suggests that even a relatively sophisticated policy-maker is going to have trouble taking DALYs calculated using CLE and adjusting them to use ALE (or the reverse), without significant amounts of time and expert assistance, as well as access to the unaggregated data.

4. The importance of user-assessability

The distinction between (relatively) user-assessable and (relatively) fixed value choices is a crucial one, ethically. As we saw above, one of the main factors motivating minimalism is the view that value judgements should be made by policy-makers, not scientists. With user-assessable value choices, like age-weighting and discounting, this distinction is tenable. It is at least potentially reasonable for an economist to say, 'We have presented you with un-age-weighted results. If you believe health losses at different ages are of different social concern, then adjust

---

[9]As mentioned above (note 7), Murray and colleagues do supply the equations to calculate DALYs without age-weights or with a different discount rate. Though this might seem to have made Arnesen and Kapiriri's task easier, in actuality removing age-weights and changing the discount rate is mathematically extremely simple. Why didn't Murray and colleagues similarly supply equations for calculating DALYs using CLE? Again, mathematically this is extremely simple. There is no simple way, however, to rewrite the standard DALY equation to do so. In the standard DALY equation, years of life lost are calculated by subtracting age at death from ALE. Since ALE is the same for any person of the same age, no additional information is needed. CLE, however, varies based not just on the age of the victim, but also sex, location, and potentially other factors (race, economic status, other morbidities, etc.). It would thus require introducing many new variables not otherwise part of the DALY equation. Indeed, this is precisely the point of this section: switching from ALE to CLE requires lots of additional information which is not readily obtainable.

these results to reflect your values before making a policy decision.' But with fixed value choices, the same division of labour doesn't work. An economist cannot say, 'We've presented you with CLE-based results. Apply whatever conception of life expectancy you deem appropriate.' The policy-maker will be unable to adjust the resulting measure. What this means is that for values like life expectancy, the policy-maker will be forced to base her policy decisions on whatever form of the measure she is given by the economist. In making a choice about whether to use counterfactual or aspirational life expectancy, the economist is, for practical purposes, deciding which value choice will be embodied in policy decisions.

Another way to think about this is to focus on the role a scientist is supposed to play in policy-making. Many scientific bodies take their goal to be to provide relevant information to decision-makers, or to facilitate informed decision-making by policy-makers.[10] Philosophers of science have argued for this view (Elliott 2006), and it coheres with the way many scientists understand their own roles. Now, it might seem that this picture supports minimalism: if a scientist's goal is to provide information (but not to influence values), then it seems natural to think that the best way to do this is to provide information that is as value-free as possible, since it is the policy-maker who should supply the values. With user-assessable values, like those connected to age-weighting, this may be correct. Un-age-weighted DALYs can directly inform a policy-maker's decision to minimize total health loss irrespective of age. And they can also indirectly inform a policy-maker's decision to minimize health loss while giving greater weight to young adults, since the policy-maker can estimate age-weighted DALYs from un-age-weighted DALYs. But this stance doesn't work with fixed values. CLE-based DALYs can provide relevant information to a policy-maker who seeks to minimize life years lost in her population. But they can't in any clear way inform the decision of a policy-maker who wants to make decisions in a way that doesn't disadvantage groups with low age-specific life expectancy. And, conversely, ALE-based DALYs can inform the latter decision, but aren't very useful for the former.[11]

Thus, if an economist takes her goal to be to provide relevant information to policy-makers, or to promote informed decision-making by policy-makers, she can't simply default to CLE on the grounds that in representing *actual* life expectancy, CLE is value-free. Economists cannot reasonably remain agnostic or neutral

---

[10]See, for example, statements by the US Bureau of Labor Statistics (http://www.bls.gov/bls/infohome. htm), the Intergovernmental Panel on Climate Change (http://www.ipcc.ch/organization/organization. shtml), and the American Institute of Biological Sciences (http://www.aibs.org/about-aibs/).

[11]In a pair of articles, McKaughan and Elliott (2013, 2018) explore the question, closely related to the question of this paper, of whether and how scientists should 'frame' their results. They argue that a balance between objectivity and informativeness can be struck by allowing scientists to frame their results while also promoting 'backtracking' on the part of information recipients: helping readers to understand how value judgements have influenced scientists' presentation of results, and '[clarifying] how one could arrive at alternative interpretations based on different value judgments' (McKaughan and Elliott 2018: 198). If my argument is correct, enabling something like 'backtracking' may be an appropriate goal when it comes to user-assessable values, but it won't work well for fixed values. As the Arnesen and Kapiriri paper shows, in such cases it may be difficult or impossible even for other experts to 'backtrack', even if the original authors are transparent about their value judgements. If, then, McKaughan and Elliott are correct (as I think they are) that there is a tension between objectivity and informativeness, my argument shows that their solution can only be a partial one.

concerning fixed value choices. In such cases, whatever decision they make will dictate, or at least significantly constrain, the values which influence downstream policy decisions. This means that a positive choice – between CLE, ALE and whatever other options there are – must be made and is ethically loaded in a way that the choice not to age-weight or not to discount would not be.

All of this suggests the following general principle:

- If values of type A cannot easily be adjusted for after-the-fact by decision-makers in some measure, M (i.e. A-values are *fixed* for M), there is strong reason for economists to structure M to reflect a preferred position on A, even if they are uncertain about what views on A are widely held, and/or about the correct view on A.

If this principle is correct, then minimalism is false, as is Murray's version of the intermediate view. Though I haven't offered an argument for it here, I think those of us who are sympathetic towards the minimalist's preference for a division of labour between economists and policy-makers can comfortably endorse the following parallel principle:

- If values of type B can easily be adjusted for after-the-fact by decision makers in some measure, M (i.e. B-values are *user-assessable* for M), then in at least many cases[12] there is reason for economists to adopt a value-neutral position on B in M, even if they are relatively confident about what views on B are widely held, and/or about the correct view on B.

If this principle is correct, then maximalism is also false. Thus, the correct view to take on values in economic measures will be a version of the intermediate view other than Murray's.

Before reflecting on the importance of this result, let me close this section by clarifying one point and then responding to an obvious and important objection. First, the clarification: the above principles suggest that when structuring measures, economists need to (1) think about which policy-relevant values will end up being fixed (vs. user-assessable) in the resulting measure, and (2) for fixed values, need to substantively engage with them, making a positive determination. But thus far I have not said anything about *how* they should make that determination. There are many options. For example, economists could seek to determine (perhaps through empirical study) what position has plurality support among the decision-makers likely to use the study, or among the populations that will be affected by associated policy decisions. Or they could engage in substantive normative reasoning, to try to determine the position on the issue most likely to be objectively correct (or least likely to lead to rights violations, etc.). Or we could give economists free choice to adopt their own preferred position. And so forth.

---

[12]I include this caveat because I don't think that economists should refrain from structuring measures to reflect all user-assessable values. For example, if there is near unanimity about some value, or if policy-makers explicitly request that some value be built into a measure, that might qualify as decisive reason to do so, rendering user-assessability irrelevant.

Which of these approaches is appropriate is a difficult, important and probably context-sensitive issue, but in this paper I take no stand on it.[13] The point is that all of these require substantive engagement with the issue, rather than a retreat to a supposedly neutral position.

Second, the objection: there is one natural way to seek to lessen the importance of fixed vs. user-assessable values, and also to avoid having to address the issue I discussed in the previous paragraph. If users will be unable to translate between DALYs calculated using CLE and DALYs calculated using ALE, the obvious solution is to provide them with both. This could be done in a traditional way (via, for example, two sets of results in a printed paper) or in a more dynamic way (an online tool which can recalculate results based on a user's preferred values). A similar motivation often lies behind 'dashboard'-type approaches, where distinct components that might otherwise be aggregated into a summary measure are presented separately, to allow users to aggregate them in whatever way they choose.[14] Strategies like this seem to make fixed values functionally user-assessable, thus relieving economists of the burden of having to make value choices themselves.

I think there is certainly a place for providing alternate results and disaggregating components in this way. But, unfortunately, it is not a panacea. First, life expectancy is far from the only fixed value choice built in to DALYs. There are more than a dozen distinct value choices that are commonly built into DALYs, many of which are probably fixed, and plenty more that are not but have been proposed for consideration (Schroeder 2017a). And there exist many different positions on each of these issues. (In this paper, I've made it sound like CLE and ALE are two options, but in fact there are many different forms each could take, and indeed the 2010 Global Burden of Disease Study uses a different version of ALE than did earlier studies (Murray *et al.* 2012).) It would be impossible to present even a reasonable fraction of the possibilities in a static format, and an online, user-customizable web tool which aspired to comprehensiveness would be hopelessly complicated.[15] At minimum, presentations like this have to be simplified – the economist will need to fix some value choices and significantly constrain others, in order to put the data in a format which will be useful and informative to decision-makers. For those value choices, the argument of this paper applies.

Second, and perhaps more importantly, the scientist's job is not simply to present decision-makers with piles of data; it is also to analyse and interpret the results. Even if the online supplement to a scientific paper or policy report includes a spreadsheet which allows users to recalculate results under a range of different value choices, the scientist still needs to choose one (or perhaps two or three) version(s) of the results

---

[13]For a hint of how such arguments might go, see e.g. Elliott (2006) or Schroeder (2017b). I explore the topic more thoroughly in an unpublished work-in-progress.

[14]Extended Cost-Effectiveness Analysis provides an example of this strategy (Verguet *et al.* 2016).

[15]See http://climatepolicysimulator.princeton.edu for an example aiming to do something like this for a few value judgements connected to climate policy. GiveWell, a charity evaluator, has created a spreadsheet that permits individuals to vary value judgements in their cost-effectiveness model: https://www.givewell.org/how-we-work/our-criteria/cost-effectiveness/cost-effectiveness-models. The Global Burden of Disease Study has created a customizable data visualization tool (http://www.healthdata.org/data-visualization/gbd-compare). It does not focus on or permit modifications of value judgements like the ones discussed here, and it is already quite complicated.

to be the subject of her analysis and discussion. And as much as we might want decision-makers to be sophisticated and patient enough to dig into the details of a study to identify methodological choices the scientist made, or at least to scour the online appendix to identify an appropriate set of alternative results, we know that this is not always or even usually the case. For better or worse, it is a scientist's top-line or primary results that will garner the bulk of most decision-makers' attention. And so even if the Global Burden of Disease Study published DALYs using four different conceptions of life expectancy, whichever was chosen as the 'primary' calculation would, in all likelihood, be the one upon which most associated policy decisions were made. And so this choice matters.[16]

5. Extending the argument

My argument has proceeded primarily via a case study concerning value choices which structure the DALY. I chose that case study, in part, because the version of the intermediate view I was considering – according to which values should structure a measure if and only if they are widely held and substantively reasonable – came from Christopher Murray, one of the original architects of the DALY. But I don't think anything I said is specific to Murray's view, to the DALY, or even to health economics as a field. Murray's view reflects, I think, the implicit practice of many economists in a range of sub-disciplines, who are often not averse to building in uncontroversial value choices, but who refrain from building in controversial ones unless absolutely necessary. And the distinction between user-assessable and fixed value choices is certainly applicable far beyond the DALY. Any complex summary measure has the potential to include fixed value choices. By definition, information is lost in moving from raw data to a summary measure, and so any value choice which can only be applied to the raw data will run the risk of being a fixed value choice. For these reasons, I think the principle I proposed in the previous section, according to which there is strong reason for an economist to structure measures to reflect preferred positions on fixed value choices (even if there is great disagreement about that value, or significant uncertainty about the correct view on that value), applies to economic measures widely.[17]

Nothing I've said, however, suggests that Murray's criteria were irrelevant – just that they were sometimes outweighed by a third factor. Indeed, I agree with Murray that, all else equal, that a value choice is widely held should count in favour of incorporating it into economic measures, and that a value can be supported with good normative reasoning also counts in favour of incorporating it into economic measures. Further, nothing I've said implies that the principle I proposed, combined

---

[16]Doesn't this also suggest that, even with user-assessable values, economists should be inclined to build a preferred position on them into the primary version of associated measures? To some extent, yes. I think economists should be more inclined than they currently are to structure measures to reflect values. However, I think there are also considerations that can count against incorporating any values, some of which I mention below (note 18). Thus, I don't think fixedness vs. user-assessability determines whether a value should be incorporated into a measure. Rather, to the extent that a value is fixed in a measure, that constitutes a strong reason to incorporate it.

[17]Indeed, I suspect it applies more broadly than that, to measures in at least some other scientific fields. But I can't pursue that idea in this paper.

with Murray's, will yield a satisfactory version of the intermediate view. I don't think it does. I suspect there are several additional factors that economists ought to consider when deciding how to structure their measures.[18] And I have said very little about how to handle cases in which the various criteria conflict with one another. As a result, I don't take myself to have come close to answering the question posed in the title of this paper. I hope, though, to have shown that the intermediate view is more attractive than either maximalism or minimalism about values in economic measures, and that finding a plausible version of the intermediate view is a challenging task, one which will result in an account more complex and nuanced than Murray's, and which will call for changes in standard economic practices.

6. A concrete recommendation: egalitarian values in summary measures of health

Fortunately, however, I don't think we need to wait until we have found a fully satisfactory version of the intermediate view to make some concrete recommendations. So let me conclude this paper by offering one. There is widespread agreement that when it comes to health, distribution matters. Our goal should not simply be to maximize total or average population health; we should prefer a more equal distribution of health, even if that comes at some cost to the total or average. Though this preference can take many different forms – e.g. egalitarian, prioritarian or sufficientarian – call any view which in practice will tend to give additional weight to improving health among the disadvantaged an *egalitarian* view. For decades, health economists have discussed how QALYs and DALYs could be adjusted to reflect egalitarian values.[19] To date, however, it remains rare to see a major health economic study that includes a preference for equality in QALYs or DALYs.[20] What is the reason for this, given that the architects of the DALY in particular have obviously not shied away from building many value judgements into that measure? It is because inequality is complex. As has been amply demonstrated (Sen 1992; Temkin 1996), there are many different senses in which a distribution can be more or less equal, and many of these senses are plausibly of moral importance. As a result, no single way of quantifying egalitarian values has gained general acceptance. Since economists have not known how to capture egalitarian values, they have understandably refrained from doing so.[21] (In terms of Murray's version of the intermediate view: there is no specific conception of inequality which is widely accepted or uncontroversial.)

Note, however, that egalitarian values will generally qualify as fixed value choices in QALYs or DALYs. To introduce distribution-sensitivity to distribution-insensitive QALYs or DALYs will require access to the disaggregated raw data, and, depending

---

[18]For example: is the value relevant to many different uses of a measure, or only a few? Is the value already reflected in other aspects/components of the measure? Have competing or complementary studies adjusted for the value in the past?

[19]See, for example, Gakidou *et al.* (2000), Anand *et al.* (2001), Adler (2012) and Eyal *et al.* (2013).

[20]The only example I'm familiar with of a major health economic study, not specifically billed as a study of inequality, to incorporate distribution-sensitivity into its primary results is the WHO's World Health Report 2000 (World Health Organization 2000). The UN's Human Development Reports also include inequality-adjusted measures of life expectancy (though these are not quality-/disability-adjusted).

[21]For an expression of this argument, see Murray (1996: 61–63).

on one's conception of inequality, may also require additional information – e.g. the socio-economic status of each individual with an adverse health outcome. It seems unlikely that decision-makers will be able to readily estimate this information from the summary statistics.[22] This suggests, I think, that health economists should seriously consider adjusting QALY- and DALY-based analyses (including cost-effectiveness analyses) to reflect egalitarian values, despite the fact that there are many different ways of doing so, none of which has received general acceptance.[23]

How should economists go about choosing a single metric for inequality in the face of such disagreement? As above, that goes beyond the scope of this paper. But, briefly, it seems to me that a satisfactory solution will require input from a wide range of groups: economists, epidemiologists, health policy-makers, philosophers, political theorists and (especially) the public. Perhaps a working group could arrive at a consensus (or, failing that, at least a strong plurality) on a conception of inequality which they all regard as sufficiently close to their own view, to be substantially preferable to distribution-neutrality.[24] To address concerns about comparability, distribution-sensitive results could be presented alongside traditional, distribution-neutral results – though I think the above considerations suggest that the distribution-sensitive results should probably be presented as the 'primary' or 'main' ones. The status quo, according to which QALYs and DALYs are presented in a distribution-insensitive way, has likely led policy-makers to make decisions in a way that embodies distribution-insensitive values, even if the policy-makers themselves hold egalitarian views (Schroeder 2017c: 1492–1493). If, as public discourse suggests, most of us hold egalitarian values when it comes to health, this is a problem – and if egalitarian values count as fixed in QALYs and DALYs, a problem that can't reasonably be attributed to or corrected for by policy-makers alone.

Author ORCIDs. 🄳 S. Andrew Schroeder 0000-0001-7305-3283

---

[22]I suspect that many of the additional criteria mentioned in note 18, above, will also count in favour of incorporating egalitarian values into DALYs. For example, egalitarian values are already built into the DALY's conception of life expectancy (discussed above), and several other aspects of the calculation (Voigt 2012; Schroeder 2017a).

[23]In this conclusion I agree with the proponents of Extended Cost Effectiveness Analysis (ECEA), who recommend that data for health policy assessments be presented in a disaggregated form which includes some distributional information (Verguet *et al.* 2016). I suspect, however, that even if ECEA caught on, there would still be a demand that it be supplemented by a single summary statistic, e.g. so that a large number of interventions could be quickly compared or ordered on league tables. My argument here, then, would suggest that this summary statistic be distribution-sensitive.

[24]The WHO's Consultative Group on Equity and Universal Health Coverage, a group including philosophers, economists, policy experts and clinicians from a range of countries, settled on a conception of equality which gives (non-absolute) priority to the worse off, in particular to those who are worse off with respect to lifetime health (World Health Organization 2014: 15). This provides some evidence that the consensus I suggest could be found, and that perhaps it would be built around a prioritarian approach.

References

Adler M. 2012. *Well-Being and Fair Distribution*. Oxford: Oxford University Press.

Anand S. and K. Hanson 1997. Disability-adjusted life years: a critical review. *Journal of Health Economics* **16**, 685–702.

Anand S. and K. Hanson 1998. DALYs: efficiency versus equity. *World Development* **26**, 307–310.

Anand S., F. Diderichsen, T. Evans, V. M. Shkolnikov and M. Wirth 2001. Measuring disparities in health: methods and indicators. In T. Evans, M. Whitehead, F. Diderichsen, A. Bhuiya and M. Wirth (eds), *Challenging Inequities in Health: From Ethics to Action*. Oxford: Oxford University Press, pp. 49–67.

Arnesen T. and L. Kapiriri 2004. Can the value choices in DALYs influence global priority-setting? *Health Policy* **70**, 137–149.

Barendregt J. J., L. Bonneux and P. J. Van der Maas 1996. DALYs: the age-weights on balance. *Bulletin of the World Health Organization* **74**, 439–443.

Bognar G. 2008. Age-weighting. *Economics and Philosophy* **24**, 167–189.

Douglas H. 2009. *Science, Policy, and the Value-Free Ideal*. Pittsburgh, PA: University of Pittsburgh Press.

Elliott K. 2006. An ethics of expertise based on informed consent. *Science and Engineering Ethics* **12**, 637–661.

Eyal N., S. A. Hurst, O. F. Norheim and D. Wikler (eds) 2013. *Inequalities in Health: Concepts, Measures, and Ethics*. Oxford: Oxford University Press.

Gakidou E., C. J. Murray and J. Frenk 2000. Defining and measuring health inequality: an approach based on the distribution of health expectancy. *Bulletin of the World Health Organization* **78**, 42–54.

Lyttkens C. H. 2003. Time to disable DALYs? *European Journal of Health Economics* **4**, 195–202.

McKaughan D. J. and K. C. Elliott 2013. Backtracking and the ethics of framing: lessons from voles and vasopressin. *Accountability in Research* **20**, 206–226.

McKaughan D. J. and K. C. Elliott 2018. Just the facts or expert opinion? The backtracking approach to socially responsible science communication. In S. Priest, J. Goodwin and M. F. Dahlstrom (eds), *Ethics and Practice in Science Communication*. Chicago, IL: University of Chicago Press, pp. 197–213.

Menzel P., M. Gold, E. Nord, J. Pinto-Prades, J. Richardson and P. Ubel 1999. Toward a broader view of values in cost-effectiveness analysis of health. *Hastings Center Report* **29**, 7–15.

Murray C. 1996. Rethinking DALYs. In C. Murray and A. Lopez (eds), *The Global Burden of Disease*. Geneva: World Health Organization, pp. 1–98.

Murray C. and A. Acharya 1997. Understanding DALYs. *Journal of Health Economics* **16**, 703–730.

Murray C., et al. 2012. Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet* **380**, 2197–2223.

Nord E. 1999. *Cost-Value Analysis in Health Care*. Cambridge: Cambridge University Press.

Polinder S., J. A. Haagsma, C. Stein and A. H. Havelaar 2012. Systematic review of general burden of disease studies using disability-adjusted life years. *Population Health Metrics* **10**, 21.

Reiss J. 2013. *Philosophy of Economics: A Contemporary Introduction*. London: Routledge.

Reiss J. 2017. Fact-value entanglement in positive economics. *Journal of Economic Methodology* **24**, 134–149.

Schokkaert E. 2015. How to introduce more (or better) ethical arguments into HTA? *International Journal of Technology Assessment in Health Care* **31**, 1–2.

Schroeder S. A. 2017a. Value choices in summary measures of population health. *Public Health Ethics* **10**, 176–187.

Schroeder S. A. 2017b. Using democratic values in science: an objection and (partial) response. *Philosophy of Science* **84**, 1044–1054.

Schroeder S. A. 2017c. Consequentializing and its consequences. *Philosophical Studies* **174**, 1475–1497.

Sen A. 1992. *Inequality Reexamined*. Cambridge, MA: Harvard University Press.

Stiglitz J. E., A. Sen and J.-P. Fitoussi 2010. *Mis-measuring Our Lives: Why GDP Doesn't Add Up*. New York, NY: The New Press.

Temkin L. 1996. *Inequality*. Oxford: Oxford University Press.

Tsuchiya A. 2000. QALYs and ageism: philosophical theories and age weighting. *Health Economics* **9**, 57–68.

Verguet S., J. J. Kim and D. T. Jamison 2016. Extended cost-effectiveness analysis for health policy assessment: a tutorial. *PharmacoEconomics* **34**, 913–923.

Voigt K. 2012. Measuring global health. In P. Lenard and C. Straehle (eds), *Global Justice and Health Inequalities*. Edinburgh: Edinburgh University Press, pp. 139–156.

Vos T. 2002. Shifting the goalpost– normative survivorship goals in health gap measures. In C. Murray, J. Salomon, C. Mathers and A. Lopez (eds), *Summary Measures of Population Health: Concepts, Ethics, Measurement and Applications*. Geneva: World Health Organization, pp. 267–272.

World Health Organization. 2000. *The World Health Report 2000: Health Systems: Improving Performance*. Geneva: World Health Organization.

World Health Organization. 2014. *Making Fair Choices on the Path to Universal Health Coverage*. Geneva: World Health Organization.

Williams A. 1997. Intergenerational equity: an exploration of the 'fair innings' argument. *Health Economics* **6**, 117–132.

Williams A. 1999. Calculating the global burden of disease: time for a strategic reappraisal? *Health Economics* **8**, 1–8.

S. Andrew Schroeder is an Associate Professor of Philosophy at Claremont McKenna College. Much of his research lies at the intersection of ethics, political philosophy and the philosophy of science. He is currently writing a series of papers exploring the ethical norms that ought to guide scientists in choosing between alternative representations of their results.