

DUKE ENVIRONMENTAL AND ENERGY ECONOMICS WORKING PAPER SERIES  
organized by the  
NICHOLAS INSTITUTE FOR ENVIRONMENTAL POLICY SOLUTIONS  
and the  
DUKE UNIVERSITY ENERGY INITIATIVE

# COST-BENEFIT ANALYSIS AND DISTRIBUTIONAL WEIGHTS: AN OVERVIEW

Matthew D. Adler\*

Working Paper EE 13-04  
August 2013

\*Duke Law School

## Acknowledgements

The author thanks James Hammitt, Alex Pfaff, and Nicolas Treich for their comments on this paper.

*The Duke Environmental and Energy Economics Working Paper Series provides a forum for Duke faculty working in environmental, resource, and energy economics to disseminate their research. These working papers have not necessarily undergone peer review at the time of posting.*



## **Introduction**

Cost-benefit analysis (CBA) is notoriously insensitive to distributional concerns. CBA quantifies well-being impacts by summing monetary equivalents: the amounts that individuals are willing to pay for policies they prefer, or to accept in return for policies they disprefer. CBA will favor a policy with a positive sum of monetary equivalents, even if some are made worse off by the policy. Nor is CBA sensitive to the distribution of these valuations across the population. For example, if a pollution-reduction policy produces net benefits for higher income individuals (who are willing to pay a lot for a cleaner environment), and net costs for lower-income consumers or workers (who must pay higher product prices or receive lower wages as a result of the policy, and would prefer on balance not to make those expenditures in exchange for the pollution reduction), CBA will choose the policy as long as the (positive) sum of monetary equivalents of the higher-income group is larger in magnitude than the (negative) sum of monetary equivalents of the lower-income individuals.

The distributional insensitivity of CBA is sometimes mitigated, in practice, by monetizing various goods and bads using constant values that do not vary with individuals' attributes—for example, population-average values. The U.S. government employs a constant “value of statistical life” (VSL)—the conversion factor establishing individuals' monetary equivalents for small changes in their fatality risks—rather than assigning a higher VSL to richer individuals, as textbook CBA would recommend. (Viscusi 2010, p. 2) Textbook CBA, with heterogeneous VSL, would recommend citing hazardous sites in poorer rather than richer neighborhoods—even if those living near the sites are not compensated. CBA with a constant VSL does not have this upshot. However, the use of population-average or otherwise constant values lacks any theoretical basis. Moreover, this practice can have unpleasant implications. For example, if a population-average VSL is used to establish mandatory safety standards for some consumer product, poorer consumers may be forced to spend more money on risk reduction than they would prefer. (Sunstein 2004).

Arguably, distributional considerations should be incorporated into CBA via so-called “distributional weights.” Monetary equivalents would be adjusted by weighting factors reflecting individuals' incomes (with lower-income individuals tending to get larger weights), or other welfare-relevant attributes such as health, life expectancy, or environmental quality.

Distributional weighting actually has a substantial pedigree. This is hardly a new idea. A scholarly literature dating from the 1950s endorses the use of weights, and analyzes how to specify them. (Boadway and Bruce 1984, pp. 271-81; Brent 1984; Cowell and Gardiner 1999; Creedy 2006; Dasgupta and Pearce 1972; Dasgupta, Sen and Marglin 1972; Dreze 1998; Dreze and Stern 1987; Fleurbaey, Luchini, Muller, and Schokkaert 2013; Johansson-Stenman 2005; Liu, 2006; Little and Mirrlees 1974; Meade 1955, ch. 2; Ray 1984; Squire and van der Tak 1975; Weisbrod 1968; Yitzhaki 2003.)

Distributional weights were adopted, for a time, at the World Bank. (Little and Mirrlees 1994). They are currently recommended by the official CBA guidance document for the U.K. (HM Treasury 2003, pp. 91-94). Distributional weights, however, appear to have been rarely if ever used by CBA practitioners in the U.S. government, and the parallel U.S. guidance document does not discuss them. (Office of Management and Budget 2003).

This Article provides an accessible introduction to the topic of distributional weights. In an age characterized by increasing income inequality, and increasing concern about inequality, it is important to consider how CBA—now the most widespread policy-evaluation tool—might be attuned to equity. The fulcrum for my discussion will be the concept of the “social welfare function” (SWF). The SWF is a foundational concept in much of welfare economics, providing the basis for optimal tax theory, growth theory, and the economic analysis of climate change. CBA with distributional weights, in turn, is a practicable method for implementing an SWF.

In fact, to the extent that scholars favor distributional weights, they generally take the perspective adopted here: CBA with weights is seen as a proxy for an SWF. This account of CBA is quite different from the well-known view which sees CBA as a tool for implementing the criterion of Kaldor-Hicks efficiency (potential Pareto-superiority). The Kaldor-Hicks criterion has the advantage of avoiding interpersonal comparability, but has various flaws, described in a literature beginning with Scitovsky (1941; see Gorman 1955; Chipman and Moore 1978; Sen 1979; Boadway and Bruce 1984). The debate about the Kaldor-Hicks criterion is surely well known to readers of this journal, and will not be recapitulated here. Less familiar, perhaps, is the existence of an *alternative* defense of CBA: a defense that rejects Kaldor-Hicks efficiency, relies instead on the notion of an SWF, and sees weighted CBA as a methodology for operationalizing an SWF. It is *this* view of CBA which is embodied in the scholarship on distributional weights that I have just mentioned, and which I will be setting forth here.

This Article is certainly *not* a comprehensive survey of this body of writing. Rather, the aim is to present a clear and readable account of how distributional weights might be specified—showing that the specification of weights, although surely value-laden, can be given intellectual structure and rigor. At the same time, the Article will address important questions that can be raised concerning the appropriateness of weights.

I first describe the SWF concept, with a particular focus on two specific SWFs: the utilitarian SWF, and an isoelastic/Atkinson SWF. The Article then discusses the functional form of weights matching these two SWFs: utilitarian weights and isoelastic weights. Next, the Article provides a concrete example (involving risk-reduction policies and VSL) to show how distributional weights might be brought to bear on environmental policy choice and, perhaps, solve dilemmas with respect to the choice between differentiated and population-average valuations of goods.

The Article concludes by considering two clusters of objections to distributional weights. One concerns heterogeneous preferences. SWFs (and thus distributional weights) require interpersonal comparisons of well-being; but how are such comparisons possible if individuals do not have the same preferences? A different cluster concerns the tax system. Aren't distributional considerations best handled by the tax system, and not distributional weights?

A third objection to weighting, not further discussed below, is that its implementation requires information about the incidence of costs and benefits across different population subgroups, and not merely population aggregate or average values. In this era of “big data” and highly sophisticated computer models of the economy and of the physical environment, it is hard to see how the third objection has much force.<sup>1</sup> Surely cost-benefit analysts otherwise persuaded of the appropriateness of weights could make *some* predictions (more or less detailed) about distributional incidence, and feed this information into the weighting scheme. (On incidence analysis of environmental policy, see Fullerton 2009; Parry et al. 2006).

The concept of weights, and challenges to weighting, are generally discussed using a simple, one-period model. This suffices to illustrate many key ideas and objections; but there are additional aspects to the topic of weighting that arise in an intertemporal context. In particular, what is the connection between distributional weights and the discount rate? The reader familiar with scholarship on climate change will note that distributional weights in the one-period case adjust for the marginal utility of consumption; and that the discount rate applied to the consumption of future generations, by climate change scholars, does the very same thing. (Stern 2007, ch. 2A; Nordhaus 2008; Dasgupta 2008.) In short, the discount rate functions as a kind of distributional weight. Unfortunately, given space constraints, this observation cannot be further developed. How to extend the model of distributional weights presented here to the intertemporal case is a topic that I must leave aside.

---

<sup>1</sup> In an influential critique of distributional weights, Harberger (1978) makes a related argument—namely, that distributional incidence depends upon demand and supply elasticities, and thus that optimal policies using weights may be counterintuitive (for example, an income tax schedule with regressive rates). The proponent of weighting can easily “bite this bullet.” Since Bentham, those working in the welfarist tradition have understood that their commitment is to a particular *criterion* for evaluating policies (for example, a utilitarian or isoelastic SWF), and not to particular *policies*. Harberger, further, argues that distributional weights may lead to policies with large efficiency costs. But efficiency costs are just costs in light of unweighted cost-benefit analysis—a methodology that, from the perspective of the SWF tradition, is problematic, since it does not correspond to any plausible SWF.

The analysis aims to be accessible. Formally inclined readers should consult the Appendix, where various claims are stated more rigorously, or proved.

## Social Welfare Functions

The concept of the social welfare function originates with work by Abram Bergson and Paul Samuelson. (Bergson 1938, 1948, 1954; Samuelson, 1947). It was reenergized by Amartya Sen, in response to Arrow’s impossibility theorem (Sen 1970), and was the basis for James Mirrlees’ pathbreaking scholarship on optimal tax theory (see Tuomala 1990). It now permeates many subdisciplines within normative economics (although less so governmental practice). See Adler (2012, pp. 79-88, summarizing scholarly development of the SWF concept).

The SWF framework has two key elements: an interpersonally comparable utility function, which transforms any given outcome (a possible consequence of policy choice) into a list or “vector” of utility numbers, one for each person in the population; and some rule for ranking these vectors.

To illustrate, imagine that the model which the policymaker uses to think about her choices characterizes each individual in terms of his consumption (an individual’s expenditure on marketed goods and services), health, and leisure. For simplicity, there are three people in the population (Jim, Sue, and Laura), and two outcomes being compared ( $x$  and  $y$ ). Jim has a particular bundle of consumption, health, and leisure attributes in  $x$ , and a different bundle in  $y$ . The same is true of Sue. Laura, as it happens is unaffected by the choice between the outcomes; her attributes are the same in both.

Our utility function assigns Jim’s bundles of attributes in  $x$  and  $y$  the utility values 10 and 11, respectively; it assigns Sue’s bundles the values 30 and 25, respectively; and Laura’s two bundles are assigned the same utility value, 40. Thus outcome  $x$  is mapped onto the utility vector (10, 30, 40), and  $y$  the vector (11, 25, 40), as shown in Table 1 immediately below.

**Table 1: Outcomes as Utility Vectors**

					
<u>Outcomes</u>			<u>Outcomes</u>		
$x$ $y$			$x$ $y$		
Jim	$\mathbf{B}_{\text{Jim}}^x$	$\mathbf{B}_{\text{Jim}}^y$	Jim	10	11
Sue	$\mathbf{B}_{\text{Sue}}^x$	$\mathbf{B}_{\text{Sue}}^y$	Sue	30	25
Laura	$\mathbf{B}_{\text{Laura}}^x$	$\mathbf{B}_{\text{Laura}}^y$	Laura	40	40

Note: In the left table,  $\mathbf{B}_{\text{Jim}}^x$  denotes the bundle of attributes (consumption, health, leisure) of individual Jim in outcome  $x$ , and so forth. In the right table, the numbers are the utilities corresponding to these attribute bundles.

Let me postpone, for the moment, the question of where these utility numbers come from; and consider different possible rules for comparing utility vectors. (See Adler 2012, ch. 5; Bossert and Weymark 2004). One such rule is the *leximin* rule: it compares the utility levels of the worst-off individuals; if those are equal the second-worst-off; and so forth. Leximin here prefers  $y$ , since Jim (the worst off) has utility 11 instead of 10.

A different rule is the utilitarian rule, which sums utilities. Utilitarianism prefers  $x$ , since the sum of utilities is 80 rather than 76.

Although utilitarianism is sensitive to the distribution of *consumption* (given the declining marginal utility of consumption), it does not take account of the distribution of utility itself. Imagine that Jim's and Laura's utilities remain as in Table 1, but that Sue's utility level in  $x$  is any number  $U$ , however large; while her utility level in  $y$  is  $(U - 1 - \epsilon)$ , with  $\epsilon$  an arbitrarily small positive number. Utilitarianism still prefers  $x$  to  $y$ .

Leximin *is* sensitive to the distribution of utility. Even though Sue's loss from  $y$  is greater than Jim's gain, leximin prefers  $y$  because Jim is worse off than Sue. However, leximin is absolutist, in the sense that it is willing to incur arbitrarily large utility losses for better-off individuals in order to realize a utility gain (however small) for an individual who is worse-off and would remain so after the gain. Imagine again that Jim's and Laura's utilities remain as in Table 1 and that Sue's utility level in  $x$  is any number  $U$ , however large; but now imagine that Sue ends up at level  $11 + \epsilon$  in  $y$ , with  $\epsilon$  an arbitrarily small positive number. Thus Sue's loss from  $y$  is the arbitrarily large  $U - 11 - \epsilon$ , while Jim's gain is only 1; and Sue ends up in  $y$  only  $\epsilon$  better off than Jim. Still, leximin prefers  $y$  to  $x$ .

The isoelastic/Atkinson<sup>2</sup> SWF lies in between leximin and utilitarianism. This SWF is parameterized by an inequality-aversion parameter  $\gamma$  and ranks utility vectors using the formula  $\frac{1}{(1-\gamma)}(u_1^{1-\gamma} \quad u_2^{1-\gamma} \quad \dots \quad u_N^{1-\gamma})$  or the formula  $\ln u_1 + \ln u_2 + \dots + \ln u_N$  in the special case of  $\gamma = 1$ . The  $\gamma$  parameter can take any positive value. If  $\gamma$  is zero rather than positive, this becomes utilitarianism. As  $\gamma$  becomes larger and large, the isoelastic SWF gives more and more priority to utility changes affecting those at lower utility levels. In the example in Table 1, if  $\gamma$  is less than or equal to approximately 1.6, the isoelastic SWF prefers outcome  $x$  to  $y$ . Sue stands to lose more utility moving to  $y$  (5 units) than Jim stands to gain (1 unit); and with a low value of  $\gamma$  the isoelastic SWF assigns a greater social value to her loss than to Jim's gain even though Jim is worse off than Sue in both outcomes. With  $\gamma$  greater than 1.6, the isoelastic SWF now prefers  $y$  to  $x$ . The priority given to utility changes affecting those who are worse off is now large enough

---

<sup>2</sup> Thus named because of its connection to the Atkinson inequality measure.

that moving Jim from 10 to 11 is seen as more socially valuable than avoiding Sue's move from 30 to 25.

This Article, as mentioned, will focus on the specification of distributional weights to approximate the utilitarian and isoelastic SWFs. Both are quite popular in the SWF literature, and the reason for their popularity can be seen axiomatically. They satisfy the Pareto principle: if at least one person's utility increases, and no one else's decreases, so does the value of the SWF. These SWFs also satisfy an axiom of anonymity/impartiality: they are indifferent between any given utility vector and all permutations (rearrangements) of its component utility numbers. In other words, anonymous/impartial SWFs focus *only* on the pattern of well-being, and not on the identities of the people who end up at particular well-being levels.

Moreover, the utilitarian and isoelastic SWFs are separable—meaning that the ranking of outcomes is not influenced by the utility levels of unaffected people. In the above example, Laura is unaffected. She happens to be at level 40 in both outcomes; but note that the utilitarian SWF would prefer  $x$  to  $y$  in *any* case where Jim's and Sue's utilities are as in Table 1 and Laura has the same utility level in the two outcomes, *regardless of what that level is*. Similarly, the isoelastic SWF with  $\gamma$  less than or equal to 1.6 would prefer  $x$  to  $y$  regardless of Laura's level, and the isoelastic SWF with  $\gamma$  greater than 1.6 would prefer  $y$  to  $x$  regardless of Laura's level. Separability occurs with these SWFs because they use an additive formula. Separability is both normatively defensible, and a huge practical advantage in policy analysis—enabling the analyst to focus her efforts on characterizing a policy's impacts on those whose well-being would be changed by the policy (and not *also* to worry about how that change would affect their position relative to the potentially vast number of unaffected).

The leximin SWF is also Paretian, anonymous, and separable, and is also popular among some SWF theorists. But it cannot be represented by a mathematical formula, which creates difficulties in mimicking this SWF with distributional weights. Moreover, using an isoelastic SWF with larger and larger values of  $\gamma$  yields a ranking of utility vectors which is closer and closer to the leximin ranking.

All SWFs require some degree of *interpersonal comparability of well-being*, in the following sense. If we start with a particular rule for ranking utility vectors (be it the utilitarian rule, the isoelastic rule, the leximin rule, or any other), and then transform the utility vectors associated with outcomes so that intrapersonal comparisons of utility levels, differences, and ratios are preserved but interpersonal comparisons are not, the ranking of outcomes may change as well.<sup>3</sup> Given a particular such rule, we won't have a stable ranking of outcomes absent some

---

<sup>3</sup> It should be noted that the isoelastic SWF with the special value  $\gamma = 1$ , i.e., the sum of the logarithm of utilities, is invariant to individual-specific ratio transformations of utility (given a fixed population size). Such transformations do preserve a kind of interpersonal comparability, namely whether the individual is better or worse off than the common zero point. See Bossert and Weymark 2004, pp. 1146-49. The utilitarian SWF, and isoelastic SWFs with other values of  $\gamma$ , are not invariant to individual-specific ratio transformations of utility.

methodology for assigning utility numbers to individuals (as a function of their attributes in outcomes) that captures not only intrapersonal well-being information, but also interpersonal well-being levels, differences, and/or ratios. (Bossert and Weymark 2004).<sup>4</sup>

More specifically, the utilitarian SWF requires interpersonal comparability of well-being differences. Imagine that two outcomes are assigned utility numbers, so that each individual’s utility difference between the outcomes has a particular magnitude. Imagine that these utility numbers are changed, so that the relative size of these differences, across individuals, is no longer the same. Then the ranking of the outcomes by the utilitarian SWF may change as well.

Consider Table 2 below, which shows possible renumberings of Jim and Sue’s utility. In the first renumbering, we rescale Jim’s utility by a Jim-specific ratio transformation (multiplying each utility value assigned to Jim by a positive number  $a_{\text{Jim}}$ ), and we rescale Sue’s by a Sue-specific ratio transformation (multiplying each value assigned to Sue by a positive number  $a_{\text{Sue}}$ ). In the second renumbering, we rescale Jim’s and Sue’s utility by a common linear transformation (multiplying all utilities by a single positive number  $a$  and adding a common  $b$ ). In the third renumbering, we rescale Jim’s and Sue’s utility by a common ratio transformation (multiplying all by a single positive number  $a$ ).

**Table 2: Interpersonal Comparisons and the Renumbering of Utility**

	<i>Original</i>			<i>Renumbering 1</i>			<i>Renumbering 2</i>			<i>Renumbering 3</i>		
	<i>x</i>	<i>y</i>	<b>Diff.</b>	<i>x</i>	<i>y</i>	<b>Diff.</b>	<i>x</i>	<i>y</i>	<b>Diff.</b>	<i>x</i>	<i>y</i>	<b>Diff.</b>
Jim	10	11	<b>1</b>	200	220	<b>20</b>	2	12	<b>10</b>	30	33	<b>3</b>
Sue	30	25	<b>-5</b>	60	50	<b>-10</b>	202	152	<b>-50</b>	90	75	<b>-15</b>
Sum	40	36		260	270		204	164		120	108	

Note: In the first renumbering, Jim’s utilities are each multiplied by a Jim-specific factor  $a_{\text{Jim}}$  equaling 20, while Sue’s are multiplied by a Sue-specific  $a_{\text{Sue}}$  equaling 2. In the second renumbering, Jim’s and Sue’s utilities are subject to a common linear transformation  $au + b$ , with the common scaling factor  $a$  equaling 10 and the constant  $b$  equaling  $-98$ . Finally, in the third renumbering, Jim’s and Sue’s utilities are subject to a common ratio transformation, each multiplied by the common positive number 3.

The bold column labeled “**Diff**” shows the difference between each person’s utility in  $y$  and  $x$ . Note that the relative size of these differences (Sue’s difference divided by Jim’s) is  $-5$ , and that the second and third renumbering preserves this relative size while the first renumbering does not. The last row displays the sum of utilities.

<sup>4</sup> A different possible understanding of the “social welfare function” employs a specific rule for ranking outcomes as a direct function of individual attributes, and then mirrors this ranking using a rule for ranking utility vectors that varies along with the assignment of individual utilities. (Fleurbaey, Luchini, Muller and Schokkaert 2013). This approach is not considered here.

With the original scheme of utility assignments, Sue's utility difference between the outcomes is  $(-5)$  and Jim's is 1. The first renumbering changes the relative magnitude of these differences. By contrast, the second and third renumberings do not change the relative magnitude of these differences. And now observe that the first renumbering alters the utilitarian ranking of the outcomes, while the second and third renumberings do not.

The isoelastic SWF is *more* demanding in terms of interpersonal comparability than the utilitarian SWF. While the utilitarian SWF requires interpersonal comparability of differences, the isoelastic SWF requires interpersonal comparability of levels, differences, and ratios. Note that the second renumbering (rescaling utilities by a common linear transformation) changes well-being ratios. (The ratio between Sue and Jim in  $x$  is originally 3 to 1, while with the second renumbering it becomes 101 to 1. The ratios in  $y$  also change.) Given a particular value of inequality aversion  $\gamma$ , the isoelastic SWF might prefer  $x$  to  $y$  using the original numbering but not the second renumbering. However, the third renumbering (multiplying utilities by a common ratio transformation) preserves well-being ratios. If the isoelastic SWF with a particular value of  $\gamma$  prefers one outcome over the other using the original numbering, then it does so with the third renumbering.

It should also be noted that the isoelastic SWF, unlike the utilitarian SWF, requires utility numbers to be positive. For a further discussion, see Appendix.

We can now address two related questions. First, what is the normative basis for a determination that one SWF is "better" than another, or more generally for a determination that the SWF framework for social decision is "better" than alternative frameworks? Second, what is the basis for assigning utility numbers that are interpersonally comparable?

On the view taken here, the SWF is a template for *ethical/moral preferences*. (See Bergson 1948, 1954; Samuelson 1947, p. 221; Harsanyi 1977, ch. 4). The term "ethical" is more common among economists, "moral" among philosophers, but the two terms are for all intents and purposes synonyms—denoting a standpoint of impartiality, whereby the decisionmaker gives equal weight to everyone's interests (or at least the interests of everyone within some population). An SWF constitutes a systematic, coherent framework for structuring ethical/moral preferences: a framework which a decisionmaker who has adopted the standpoint of impartiality might wish to use in refining and specifying her ethical/moral tastes.

Is there some deeper criterion of moral truth that establishes whether someone's ethical/moral preferences are "correct" or "incorrect"? That is a question of metaethics which is debated by philosophers, but is beyond the scope of welfare economics and, in any event, not relevant to the discussion here. *Whatever* the nature of ultimate moral truth, a decisionmaker motivated by impartial concern will need to figure out what her ethical/moral preferences are. The SWF framework is a plausible format for regimenting those preferences: a format that

conforms to various axioms which seem (to those in the SWF tradition) morally very attractive, and which the decisionmaker may find attractive as well.

While the SWF framework itself is a tool for specifying and refining the ethical/moral preferences of some decisionmaker, the *inputs* for the framework are utility numbers measuring the well-being of everyone within some population (relative to which the decisionmaker has adopted an attitude of impartiality). Very plausibly, there is a close connection between someone's welfare and that individual's *personal*, i.e., self-interested, preferences. Phillip is better off in outcome  $x$  than  $y$  if, and only if, Phillip has a personal preference for  $x$ . (Adler 2012, ch. 3; Fleurbaey, Luchini, Muller, and Schokkaert, 2013).

Moreover, if members of the population have *identical* personal preferences, then interpersonal comparisons become straightforward. Joe with one bundle of attributes is better off than Raj with another if everyone (given their common personal preferences) prefers Joe's bundle to Raj's. Similarly, interpersonal comparisons of well-being differences and ratios become straightforward given homogeneous personal preferences. My analysis of distributional weights will start with this simplifying assumption of identical personal preferences—postponing, until later in the Article, the crucial (but contested) question of how the SWF framework should handle heterogeneity of personal preferences.

To reiterate: the SWF approach offers a systematic and coherent framework for regimenting the ethical/moral preferences of some decisionmaker, e.g., an elected official with power to set policy direction. And the utility numbers in the SWF represent the personal (self-interested) preferences of the affected population—which, in the simplest case, are identical. Although it would be an exaggeration to say that this way of viewing the SWF framework is the *only* perspective taken in the academic literature, it certainly has substantial support therein. In turn, this perspective will be very useful in seeing how to specify distributional weights, and in understanding the debate about weights.<sup>5</sup>

---

<sup>5</sup> This perspective also sheds light on the use of empirical data to calibrate an SWF and, in particular, to specify distributional weights. Because the decisionmaker has moral preferences that make reference to the personal preferences of individuals affected by her (the decisionmaker's) choices—because the SWF takes utilities measuring the population's personal preferences as its inputs—it seems perfectly sensible for the decisionmaker to engage in empirical research to help determine the structure of those personal preferences. (For example, she might engage in empirical research to ascertain the  $\lambda$  parameter for the CRRA utility function; see below.)

By contrast, empirical research to determine the functional form of the SWF itself—for example, to specify the  $\gamma$  parameter for the isoelastic SWF—is more puzzling. (On such research, see Lambert, Millimet, and Slottje 2003). How does such research help the decisionmaker determine what her own moral preferences are? (Admittedly, an elected decisionmaker might have a strategic interest in having moral preferences that do not deviate too far from the median voter's.)

## Utilitarian Distributional Weights

A simple one-period model will illustrate the workings of CBA and the specification of utilitarian distributional weights. In any given outcome, each individual faces a particular set of market prices, and is able to expend a certain amount (her consumption) on marketed goods and services. She also has non-consumption attributes, such as health, leisure or environmental quality. (In general, an individual's income need not equal her consumption; but in a one-period model without inheritance or bequests they are identical, and so the reader in this discussion can substitute "income" for "consumption" if she likes.)

Each policy choice, including the status quo choice of inaction, leads for sure to some outcome. Let  $\mathbf{p}^s$  denote the prices in the status quo and  $\mathbf{p}^x$  the prices in outcome  $x$ .  $c_i^s$  denotes the total consumption of individual  $i$  in the status quo, and similarly  $c_i^x$  denotes her total consumption in a given outcome  $x$ . Finally,  $\mathbf{a}_i^s$  and  $\mathbf{a}_i^x$  denote the non-consumption attributes of individual  $i$  in, respectively, the status quo and outcome  $x$ . We can use the compact notation  $\mathbf{B}_i^x$  or  $\mathbf{B}_i^s$  to denote someone's overall bundle in a given outcome, i.e.,  $\mathbf{B}_i^x = (c_i^x, \mathbf{p}^x, \mathbf{a}_i^x)$ . There are  $N$  individuals total in the population.<sup>6</sup>

Unweighted CBA sums monetary equivalents. These can be defined as "equivalent variations" (changes to status quo consumption) or "compensating variations." (Freeman 2003, ch. 3). Individual  $i$ 's equivalent variation for outcome  $x$  is the amount  $\Delta c_i^x$  such that she is indifferent between the bundle  $(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s)$  and the bundle  $(c_i^x, \mathbf{p}^x, \mathbf{a}_i^x)$ . Individual  $i$ 's compensating variation for outcome  $x$  is the amount  $CV_i^x$  such that she is indifferent between the bundle  $(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)$  and the bundle  $(c_i^x - CV_i^x, \mathbf{p}^x, \mathbf{a}_i^x)$ .

Equivalent variations are theoretically preferable to compensating variations. CBA with compensating variations can violate the Pareto principle, while CBA with equivalent variations cannot. (See Appendix.) My presentation will henceforth focus on the equivalent variation—now using the term "monetary equivalent" to mean specifically that. In practice, CBA analysts often employ compensating variations, which can be seen as a rough-and-ready proxy for the equivalent variation.

Consider now the utilitarian SWF. Recall that—in order to arrive at interpersonally comparable utilities—we are starting with the simplified (and concededly quite unrealistic) assumption of common personal preferences. Moreover, utilitarianism requires a *cardinal* interpersonally comparable utility function—a function that embodies information about utility differences.

Given common personal preferences, expected utility theory provides a straightforward path to cardinal and interpersonally comparable utilities. This theory shows that if someone has a

---

<sup>6</sup> How to specify the SWF approach, and CBA with distributional weights, to take account of variable population size is an important topic—but one that cannot be addressed here given space constraints.

well-behaved ranking of lotteries over attribute bundles, there will exist a so-called von-Neumann/Morgenstern (vNM) utility function  $u(\cdot)$  such that the *mathematical expectation* of  $u(\cdot)$  will correspond to this ranking. Moreover, another function  $v(\cdot)$  will be such that its mathematical expectation *also* corresponds to the ranking of lotteries if and only there is a positive number  $a$  and a number  $b$  such that  $v(\mathbf{B}) = au(\mathbf{B}) + b$  for any bundle  $\mathbf{B}$ . (Kreps, 1988).

Assume that individuals have common personal preferences (including lottery preferences), and that these are represented by function  $u(\cdot)$ . Let  $\mathbf{B}_i^x$  be individual  $i$ 's bundle of attributes in outcome  $x$ , and  $\mathbf{B}_i^y$  her bundle in outcome  $y$ . Using function  $u(\cdot)$ , the utilitarian SWF compares  $x$  and  $y$  by comparing the sum  $u(\mathbf{B}_1^x) + u(\mathbf{B}_2^x) + \dots + u(\mathbf{B}_N^x)$  to the sum  $u(\mathbf{B}_1^y) + u(\mathbf{B}_2^y) + \dots + u(\mathbf{B}_N^y)$ . Now, consider any function  $v(\cdot)$  that equally well represents the common lottery preferences. Using function  $v(\cdot)$ , the utilitarian SWF compares  $x$  and  $y$  by comparing the sum  $v(\mathbf{B}_1^x) + v(\mathbf{B}_2^x) + \dots + v(\mathbf{B}_N^x)$  to the sum  $v(\mathbf{B}_1^y) + v(\mathbf{B}_2^y) + \dots + v(\mathbf{B}_N^y)$ . But because  $v(\cdot)$

individual's consumption (total monetary expenditure). But equal money changes do not necessarily correspond to equal changes in interpersonally comparable utility. Adding \$100,000 to Bill Gates' consumption makes a much smaller difference to his utility than adding \$100,000 to the expenditures of someone with an ordinary income. A small money change, for a given individual, can be translated into a utility change by using the individual's marginal utility of money as an adjustment factor.

How do we calculate  $MU_i$ ? Function  $u(\cdot)$  captures individuals' (common) personal preferences over lotteries, where the "prizes" are bundles consisting of consumption amounts, market prices, and non-consumption attributes. Even if individual *do* have common personal preferences, estimating their ranking of such lotteries seems very complex.

However, recall that we are trying to estimate individuals' marginal utilities in the status quo—where individuals have various consumption and non-consumption attributes, but face a single price vector. We can therefore ignore the full content of  $u(\cdot)$ , and focus on estimating the  $u(\cdot)$  values of consumption-nonconsumption bundles given that prices  $\mathbf{p}$  are at the value  $\mathbf{p}^s$ .

There are many different empirical methods for estimating these  $u(\cdot)$  values. (See, e.g., Keeney and Raiffa 1993; Evans and Viscusi 1991; Rey and Rochet 2004; Finkelstein, Luttmer and Notowidigdo 2008). Given space limitations, I can only briefly describe one such approach, as follows. Imagine that we know how individuals rank lotteries over consumption amounts, given that non-market attributes are fixed at various levels ( $\mathbf{a}$ ,  $\mathbf{a}^*$ ,  $\mathbf{a}^{**}$ , ...). Each such ranking is captured by a "conditional" utility function  $h^{\mathbf{a}}(\cdot)$ ,  $h^{\mathbf{a}^*}(\cdot)$ ,  $h^{\mathbf{a}^{**}}(\cdot)$ , ..., which takes consumption alone as its argument. A given  $h^{\mathbf{a}}(\cdot)$  will be concave in consumption if individuals are risk-averse over consumption lotteries holding fixed non-market attributes at the designated level  $\mathbf{a}$ ; convex in consumption if individuals are risk-seeking; and linear if individuals are risk-neutral.

Assume that we have estimated this conditional utility function for various background levels of non-market attributes  $\mathbf{a}$ ,  $\mathbf{a}^*$ ,  $\mathbf{a}^{**}$  ... . Assume, moreover, that we have ordinary (non-lottery) willingness-to-pay/willingness-to-accept data, indicating the change in consumption that suffices to compensate individuals, at various consumption levels, for a change in non-market attributes from  $\mathbf{a}$  to  $\mathbf{a}^*$ ,  $\mathbf{a}^*$  to  $\mathbf{a}^{**}$ , etc. Putting both sorts of data together, we are in a position to estimate  $u(\cdot)$ , i.e., utility as a function of *both* consumption and non-market attributes, and thus the weighting factor  $MU_i$ .

Why is the approach just sketched a useful way (although not the only way) to think about the estimation of  $MU_i$ ? First, it is cognitively less challenging for individuals to reflect about their preferences for lotteries over consumption holding fixed other attributes, as opposed to lotteries with all attributes varying. Second (and relatedly) there is a vast econometric literature that tries to estimate utility as a function of income (or wealth or consumption) alone. (See Adler 2012, pp. 291-94, providing references.) Implicitly if not explicitly, what this literature is doing is inferring preferences for consumption (or income or wealth) gambles,

holding fixed other attributes. Finally, willingness-to-pay/accept data for non-market goods is also plentiful—indeed, inferring such amounts is essential for ordinary CBA even without weights, and is therefore a key research question in applied economics.

Note that the conditional function,  $h^a(\cdot)$ , is indexed to the level of non-market attributes. This allows for the possibility that the degree to which individuals are risk-averse or -seeking with respect to the consumption gambles depends upon  $\mathbf{a}$ . But two useful simplifications can now be introduced. First, we might assume that preferences over consumption gambles are invariant to the level of non-market attributes. If so, there is a single conditional function  $h(\cdot)$ , and it can be shown that  $MU_i(\cdot)$  is just equal to  $h' (c_i^s) \times m(\mathbf{a}_i^s)$ , with  $h' (\cdot)$  the first derivative of  $h(\cdot)$  and  $m(\cdot)$  a multiplicative scaling factor that scales up or down this marginal  $h(\cdot)$  value to take account of non-market attributes.

A second simplification is to assume that preferences over consumption gambles, holding fixed non-market attributes, take the constant-relative-risk-aversion (CRRA) form:

$h(c) = \frac{1}{1-\lambda} c^{1-\lambda}$  or  $\ln c$  with  $\lambda = 1$ . (Gollier 2001, ch. 2). The CRRA form is extremely popular in the literature on preferences for consumption (or income or wealth) gambles. It allows us to capture the degree of risk-aversion (or proneness) with respect to consumption in a single parameter  $\lambda$ .<sup>8</sup> Indeed, much of the existing work on distributional weights incorporates the CRRA form. For estimates of  $\lambda$ , see Kaplow (2005).

There is a mathematical isomorphism between the formula for CRRA utility (as a function of consumption), and the formula for the isoelastic SWF (as a function of individuals' utilities). However, the risk-aversion parameter  $\lambda$  of the CRRA utility function and the inequality-aversion parameter  $\gamma$  of the isoelastic SWF are conceptually quite different.  $\lambda$  is a number that captures individuals' *personal* preferences over consumption gambles. It is useful in estimating distributional weights both for the utilitarian SWF (which has no  $\gamma$  parameter) and for the isoelastic SWF (which does). By contrast, the inequality-aversion parameter  $\gamma$  captures the *moral* preferences of a certain kind of social planner (namely, one who has moral preferences representable via an SWF, and more specifically has a moral preference for equalizing well-being rather than simply aggregating utilities in utilitarian fashion).

The following table summarizes the three formulas for CBA with distributional weights to mimic a utilitarian SWF: the general case, and with the simplification, first, of a preferences for consumption lotteries that are invariant to background attributes; and, second, of invariance coupled with CRRA.

---

<sup>8</sup> The constant-*absolute*-risk-aversion (CARA) utility function also takes a simple (exponential) form, but it is widely seen to be implausible that absolute risk aversion would remain constant rather than decrease (Gollier 2001, ch. 2); and CARA utility is used much less often in work on distributional weights than CRRA utility.

**Table 3: CBA with Utilitarian Distributional Weights**

<b><u>General Case</u></b>	<b><u>Invariance</u></b>	<b><u>Invariance plus CRRA</u></b>
<p>A given outcome <math>x</math> is assigned the sum of <math>MU_i \times \Delta c_i^x</math>, summing across all individuals. <math>MU_i</math>, the marginal utility of consumption for individual <math>i</math> in the status quo, is some function of her status quo consumption and nonconsumption attributes.  <math>MU_i = f(c_i^s, \mathbf{a}_i^s)</math></p>	<p>Individuals' rankings of consumption lotteries, holding fixed background attributes, are invariant to the level of such attributes.</p> <p>The formula for the general case simplifies so that <math>MU_i = h'(c_i^s) \times m(\mathbf{a}_i^s)</math>, with <math>h(\cdot)</math> a utility function representing the ranking of consumption lotteries, and <math>h'(\cdot)</math> the first derivative</p>	<p>CRRA means that <math>h(\cdot)</math> takes the form: <math>h(c) = \frac{1}{1-\lambda} c^{1-\lambda}</math>.</p> <p>Parameter <math>\lambda</math> represents the degree of risk aversion.</p> <p>The formula for the general case simplifies so that <math>MU_i = (c_i^s)^{-\lambda} \times m(\mathbf{a}_i^s)</math></p>

Note: These are formulas for distributional weighting under conditions of certainty, with each policy choice leading for sure to one outcome. For corresponding formulas under uncertainty, see Appendix.

Note that, in all three cases, the weighting factor  $MU_i$  for individual  $i$  depends on *both* her status consumption  $c_i^s$ , *and* her status quo non-consumption attributes. Much work on distributional weights is yet more simplified—ignoring non-consumption attributes, and making someone's weighting factor *just* a function of his consumption (or income or wealth). (See, e.g., HM Treasury 2003.)

Distributional weights of this consumption-only form can be theoretically supported only in two special cases: (1) those affected by the policy are heterogeneous with respect to status quo consumption, but relatively homogeneous with respect to status quo non-consumption attributes; or (2) the utility function  $u(\cdot)$  not only satisfies the invariance requirement, but takes a special additively separable form—which has the upshot that the multiplicative scaling factor  $m(\cdot)$  will be unity for all non-consumption attributes. Even if neither of these theoretical justifications for consumption-only weights is applicable, data limitations may perhaps justify this approach: good information regarding the correlation of consumption and non-consumption attributes among the affected population may not be available (within her informational budget) to the decisionmaker.

The model presented here has assumed that each policy choice, for certain, leads to a particular outcome. More realistically, each policy choice is a probability distribution over outcomes (with these probabilities capturing the decisionmaker's uncertainty about which outcome will result), and the status quo itself is a probability distribution over outcomes. In this case, CBA defines an individual's equivalent variation for a policy  $P$  (denote this as  $\Delta c_i^P$ ) as the

change to her consumption in all status quo outcomes that makes her indifferent between the status quo and the policy.

The utilitarian SWF, in turn, can now be approximated by a formula that smoothly generalizes that given above: the sum of  $EMU_i \times \Delta c_i^P$ , where  $EMU_i$  is individual  $i$ 's expected marginal utility of consumption, given the probability distribution over consumption and other attributes that she faces in the status quo.

The use of weighted CBA to mimic a utilitarian SWF also generalizes to the case of “inframarginal” changes, where policies are no longer “small” variations from the status quo. The formula for assigning weights becomes more complicated. An individual’s weight, now, is policy-specific. It depends not only on her status quo attributes, but also the magnitude of the equivalent variation corresponding to a particular policy.

### **Isoelastic Distributional Weights**

Let us continue using the simple one-period model of choice under certainty introduced in the previous section. However, we now assume that the decisionmaker’s ethical/moral preferences take the form of an isoelastic SWF. For short, these are “isoelastic” moral preferences.

The isoelastic SWF’s ranking of policies that are “small” deviations from the status quo—like the utilitarian ranking—can be approximated by the sum of individual monetary equivalents, each multiplied by a weighting factor that is just a function of the individual’s status quo attributes. What is the functional form of these isoelastic distributional weights? Recall that isoelastic moral preferences give greater priority to well-being changes affecting worse-off individuals. Recall, too, that the degree of such priority is embodied in an inequality-aversion parameter  $\gamma$ , which can take any positive value. These features of the isoelastic SWF are reflected in the corresponding distributional weights. Individual  $i$ 's isoelastic distributional weight is her utilitarian distributional weight,  $MU_i$  — her marginal utility of income, given her status quo attributes — multiplied by an *additional* term,  $MMVU_i$ .

$MMVU_i$  is the *marginal moral value of utility* for individual  $i$ . With  $u(\cdot)$  the utility function representing individuals’ common personal preferences, and  $(c_i^s, \mathbf{a}_i^s)$  individual  $i$ 's bundle of consumption and non-consumption attributes in the status quo,  $MMVU_i$  is equal to  $u(c_i^s, \mathbf{a}_i^s)^{-\gamma}$ .  $MMVU_i$  measures the moral benefit produced by a small increase in individual  $i$ 's level of utility—the extent to which a small increase in individual  $i$ 's level of utility in a given outcome (here, the status quo outcome) improves the position of that outcome in the isoelastic moral ranking of outcomes. Note that  $MMVU_i$  is a decreasing function of utility level, given that  $\gamma$  is positive: If the utility level of individual  $i$  is larger than that of individual  $j$ ,  $MMVU_i$  is less than  $MMVU_j$ . This reflects the priority that the isoelastic SWF gives to worse-off

individuals. A small increment to individual  $j$ 's utility yields a larger moral benefit than the very same increase in individual  $i$ 's utility.

$MMVU_i$  is not merely decreasing with utility level, but its specific magnitude depends upon the value of  $\gamma$ . In determining a specific such value, the decisionmaker makes an ethical/moral choice. Having decided to endorse the isoelastic SWF (rather than the utilitarian SWF or some other functional form), she now makes the *further* decision to endorse a particular degree of moral priority for the worse off.

There are various thought experiments that the decisionmaker might undertake in thinking through this latter choice. (Adler 2012, pp. 392-99). One such thought experiment is a “leaky bucket” experiment. The decisionmaker asks herself: if High is at  $K$  times the level of well-being of Low, and we reduce High’s utility by a small amount  $\Delta u$ , and at the same time increase Low’s by some fraction of  $\Delta u$ , what is the minimum such fraction such that the combination of the loss to High and gain to Low is on balance a moral benefit? Another such thought experiment is an “equalization” experiment. If High is at  $K$  times the level of well-being of Low, and we equalize their well-being levels at some level less than halfway between the starting points, what is the smallest such level such that the combination of the loss to High and gain to Low is a moral benefit? The answer to either type of question fixes a value for  $\gamma$ .

Table 4 provides formulas for CBA with both utilitarian and isoelastic distributional weights. It covers the general case; the special case in which personal preferences over consumption lotteries are invariant to the level of non-income attributes; and the extra simplification that comes from coupling this invariance assumption with a CRRA function for the utility of consumption.

**Table 4: CBA with Utilitarian and Isoelastic Distributional Weights**

	<b>General Case</b>	<b>Invariance</b>	<b>Invariance plus CRRA</b>
<b>Utilitarian Weights</b>	Sum of $MU_i \times \Delta c_i^x$ $MU_i = f(c_i^s, \mathbf{a}_i^s)$	Sum of $MU_i \times \Delta c_i^x$ $MU_i = h'(c_i^s) \times m(\mathbf{a}_i^s)$	Sum of $MU_i \times \Delta c_i^x$ $MU_i = (c_i^s)^{-\lambda} \times m(\mathbf{a}_i^s)$
<b>Isoelastic Weights</b> $\gamma > 0$ is coefficient of inequality aversion	Sum of $MMVU_i \times MU_i \times \Delta c_i^x$ $MU_i = f(c_i^s, \mathbf{a}_i^s)$ $MMVU_i = u(c_i^s, \mathbf{a}_i^s)^{-\gamma}$	Sum of $MMVU_i \times MU_i \times \Delta c_i^x$ $MU_i = h'(c_i^s) \times m(\mathbf{a}_i^s)$ $MMVU_i = u(c_i^s, \mathbf{a}_i^s)^{-\gamma}$ , with $u(c_i^s, \mathbf{a}_i^s) = h(c_i^s) \times m(\mathbf{a}_i^s) + k(\mathbf{a}_i^s)$	Sum of $MMVU_i \times MU_i \times \Delta c_i^x$ $MU_i = (c_i^s)^{-\lambda} \times m(\mathbf{a}_i^s)$ $MMVU_i = u(c_i^s, \mathbf{a}_i^s)^{-\gamma}$ , with $u(c_i^s, \mathbf{a}_i^s) = h(c_i^s) \times m(\mathbf{a}_i^s) + k(\mathbf{a}_i^s)$ and $h(c) = (1-\lambda)^{-1} c^{1-\lambda}$

Note: These are formulas for distributional weighting under conditions of certainty, with each policy choice leading for sure to one outcome. For corresponding formulas under uncertainty, see Appendix.

Apart from the use of the extra  $MMVU_i$  term, there are three additional respects in which isoelastic weighting differs from utilitarian weighting. First, as mentioned earlier, the isoelastic SWF requires interpersonal comparability of utility ratios. Even on the assumption of common personal preferences, a vNM utility function representing such preferences is only unique up to a linear transformation, and thus is adequate for utilitarian weights but not isoelastic weights. To understand this point, consider Table 5 immediately below. Individuals are identical except for their consumption amounts. One CRRA function assigns utility numbers to these amounts. A second function is a linear transformation of the first: two times the first plus a non-zero constant. These two functions are equally good vNM representations of a possible preference structure with respect to consumption gambles: they imply the very same ranking of such gambles. Moreover, as shown in the table, the two functions imply equivalent utilitarian distributional weights, but not equivalent isoelastic weights. (On the importance of ratio-scale measurability for the isoelastic SWF, with reference to environmental policy, see Johansson-Stenman 2000).

**Table 5: Isoelastic Weights are Affected by Rescalings of the Utility Function that do not Preserve Utility Ratios**

Consumption ( $c_i$ )	Utilitarian weight with utility $u = \ln c$ . $MU_i = 1/c_i$	Utilitarian weight with utility $u^* = 2 \times \ln c + 3$ . $MU_i = 2/c_i$	Isoelastic weight, $\gamma = 2$ , with utility $u = \ln c$ $MU_i = 1/c_i$ $MMVU_i = (\ln c_i)^{-2}$	Isoelastic weight, $\gamma = 2$ , with utility $u^* = 2 \times \ln c + 3$ $MU_i = 2/c_i$ $MMVU_i = (2 \times \ln c_i + 3)^{-2}$
\$20,000	1/20000	2/20000	$(1/20000) \times 9.90^{-2}$	$(2/20000) \times 22.81^{-2}$
\$40,000	1/40000	2/40000	$(1/40000) \times 10.60^{-2}$	$(2/40000) \times 24.19^{-2}$
\$60,000	1/60000	2/60000	$(1/60000) \times 11.00^{-2}$	$(2/60000) \times 25.00^{-2}$
\$80,000	1/80000	2/80000	$(1/80000) \times 11.29^{-2}$	$(2/80000) \times 25.58^{-2}$
\$100,000	1/100000	2/100000	$(1/100000) \times 11.51^{-2}$	$(2/100000) \times 26.03^{-2}$

Note: Let  $u(c, \mathbf{a}) = \ln c \times m(\mathbf{a})$ . Let  $u^*(.) = 2u(.) + 3$ . Assume that all individuals have identical nonconsumption status quo attributes  $\mathbf{a}^s$  with  $m(\mathbf{a}^s) = 1$ . While the utilitarian weights according to  $u^*(.)$  are not numerically identical to the utilitarian weights according to  $u(.)$ , it can be seen that they are scaled up from the  $u(.)$  weights by a factor of 2, and are thus “equivalent” in the sense of producing an identical ranking of outcomes when incorporated into weighted CBA. By contrast, isoelastic weights according to  $u^*(.)$  are not scaled up from  $u(.)$  by any constant factor.

Given common personal preferences, a utility function unique up to a ratio transformation (and thus sufficient to determine isoelastic weights) can be produced by taking the vNM function and then assigning the number zero to a “threshold bundle”: a bundle of attributes which individuals regard as being just at the threshold of a life worth living (e.g., a bundle with very low consumption and bad health). This kind of “zeroing out” is familiar from

scholarship on QALYs (quality adjusted life-years), a technique for measuring health states on a zero-one scale that sets zero as a health state no better than death.

A second additional difference from utilitarian weights concerns policy choice under uncertainty. (Adler 2012, ch. 7; Fleurbaey 2010). It turns out that there are *two* ways to apply an isoelastic SWF to a set of policies, each of which is a probability distribution over outcomes: the “ex ante” approach and the “ex post” approach. The ex ante approach is readily operationalized by taking monetary equivalents under uncertainty, and by setting the weighting factor equal to the  $EMU_i$  term multiplied by a  $MMVEU_i$  term (“marginal moral value of expected utility”) analogous to the  $MMVUI_i$  term in the certainty case. The ex post approach is much more complicated to mimic with distributional weights. If one favors an isoelastic SWF applied under uncertainty in an ex post manner, there is a strong case that the computational complexity of mimicking this approach via weighted CBA (as opposed to direct implementation of the SWF) is too great to be justified. See Appendix for details.

Finally, the utilitarian SWF yields consumption-only weights, even with population heterogeneity in non-consumption attributes, if the utility function  $u(\cdot)$  takes a special additively separable form. That is not true for the isoelastic SWF. See Appendix.

### **An Illustrative Example: VSL and Weights**

This section uses the value of statistical life (VSL) to illustrate distributional weighing. VSL is the marginal rate of substitution between survival probability and income. This is the concept used by CBA to monetize the fatality-risk-reduction benefits of environmental and other policies. VSL has become hugely important in the practice of CBA by U.S. governmental agencies, especially the Environmental Protection Agency. (Cropper, Hammitt and Robinson 2011).

I use the workhouse, one-period model of VSL that is standard in the literature. (Eeckhoudt and Hammitt 2004). Each individual in the status quo earns some income, and has some probability of surviving through the end of the current period (e.g., the current year) and consuming her income; if she doesn’t survive, the income is bequested. Policies change individuals’ survival probabilities or incomes. In other words, each individual in any given *outcome* has an attribute bundle consisting of an income amount plus a single binary non-income attribute: surviving the current period, or dying. The status quo and policies are lotteries over such outcomes. Each individual has personal preferences over (income, die/survive) bundles, represented by a vNM utility function. Her VSL in the status quo reflects her status quo income and survival probability, plus these preferences.

In order to enable the calculation of distributional weights, I assume that individuals have common personal preferences over (income, die/survive) bundles, represented by a common

vNM function. I also assume that marginal utility in the death state is zero: the utility of income conditional on the attribute “die” is a flat line. Since utility here is supposed to represent *personal* benefit, this assumption seems very compelling.<sup>9</sup> It also means that we can calibrate the common vNM function by knowing the “subsistence” level of income: the level which, if combined with the attribute “survive,” is so low that individuals are indifferent between that bundle and dying. I add the simplifying assumption that individuals’ preferences over lotteries in the “survive” state are CRRA. Note finally that, by determining the subsistence level of income, we have at the very same time identified a natural zero point for purposes of the isoelastic SWF.

Given an individual’s status quo income and survival probability, we can now assign her (1) a VSL value; (2) a utilitarian weight, equaling her expected marginal utility of income ( $EMU_i$ ) and (3) an isoelastic weight, equaling  $EMU_i$  multiplied by the marginal moral value of expected utility ( $MMVEU_i$ ), with  $MMVEU_i$  in turn a function of the coefficient of inequality aversion  $\gamma$  that the policymaker chooses. For small changes in income and survival probability, an individual’s monetary equivalent (equivalent variation) is approximately the income change plus her VSL multiplied by the change in survival probability. (For example, if an individual has a VSL of \$3 million, her equivalent variation for a 1-in-1 million reduction in her fatality risk is approximately \$3, and it is approximately \$30 for a 1-in-100,000 reduction.) The utilitarian ranking of policies is, in turn, approximated by the sum of monetary equivalents multiplied by the utilitarian weights; and the ex ante isoelastic ranking by the sum of monetary equivalents multiplied by the isoelastic weights.

Assume that we are considering policies which will affect two populations: a better-off group with a higher income and lower status quo fatality risk, and a worse-off group with a lower income and higher status quo risk. Specifically, let each member of the first group have an annual income of \$100,000 and face an annual all-cause fatality risk of 0.005; while each member of the second group has an annual income of \$20,000 and faces an annual all-cause fatality risk of 0.01. (The average annual all-cause fatality risk for the entire U.S. population is in between these values, at around 0.008.) The policies will reduce fatality risks by 1 in 100,000.

Table 6 calculates VSL values for the better and worse-off individuals (for short, “Rich” and “Poor”), as well as the ratios of these values, the ratios of the Poor and Rich distributional weights, and the ratios of their isoelastic weights—given different assumptions about the coefficient of relative risk aversion  $\lambda$ , the subsistence level, and the degree of inequality aversion  $\gamma$ . Note that, especially at higher values of  $\lambda$ , the Rich VSL is many multiples of the Poor VSL. Such high multiples are inconsistent with observed values of the income elasticity of VSL—perhaps reflecting the limitations of the simple analytic model of VSL used here. (A model incorporating labor or health attributes might have less heterogeneity in VSL values.)

---

<sup>9</sup> I may have moral or altruistic reasons to care about the amount of income left to my heirs, or the state, but their consumption after my death does not change my well-being.

Alternatively, low observed income elasticities of VSL may reflect real-world violations of the axioms of expected utility. (On income elasticity of VSL, see Evans and Smith 2010; Kaplow 2005; Viscusi 2010).

**Table 6: Distributional Weights and VSL**

	$\lambda = 0.5$ Sub=1000	$\lambda = 0.5$ Sub=5000	$\lambda = 1$ Sub=1000	$\lambda = 1$ Sub=5000	$\lambda = 2$ Sub=1000	$\lambda = 2$ Sub =5000	$\lambda = 3$ Sub=1000	$\lambda = 3$ Sub =5000
VSL <sub>Rich</sub>	\$180,905	\$156,059	\$462,831	\$301,079	\$9,949,749	\$1,909,548	\$502 mill	\$20,050,251
VSL <sub>Poor</sub>	\$31,369	\$20,202	\$60,520	\$28,006	\$383,838	\$60,606	\$4,030,303	\$151,515
<b>VSL<sub>Rich</sub>/VSL<sub>Poor</sub></b>	<b>5.8</b>	<b>7.7</b>	<b>7.6</b>	<b>10.8</b>	<b>25.9</b>	<b>31.5</b>	<b>124.7</b>	<b>132.3</b>
U-Weight <sub>Poor</sub> / U-Weight <sub>Rich</sub>	2.2	2.2	5	5	24.9	24.9	124.4	124.4
Iso-Weight <sub>Poor</sub> / Iso-Weight <sub>Rich</sub> $\gamma = 0.5$	3.6	4.2	6.2	7.3	25.5	28.1	124.8	128.6
Iso-Weight <sub>Poor</sub> / Iso-Weight <sub>Rich</sub> $\gamma = 1$	5.8	7.8	7.7	10.8	26.1	31.7	125.3	133
Iso-Weight <sub>Poor</sub> / Iso-Weight <sub>Rich</sub> $\gamma = 2$	15.1	27.1	11.9	23.5	27.3	40.3	126.2	142.2

Note: “U-Weight” indicates the utilitarian distributional weight for Rich or Poor (depending on the subscript), and “Iso-Weight” the isoelastic weight. Rich individuals have an annual income of \$100,000 and fatality risk of 0.005, while Poor individuals have an annual income of \$20,000 and fatality risk of 0.01. Each column corresponds to different assumptions about the coefficient of risk aversion  $\lambda$  for CRRA utility, and the subsistence level of income.

In any event, what bears especial note about Table 6 is how weights counteract the higher VSL values of Rich individuals. For example, at  $\lambda = 2$  and a subsistence level of 1000, the Rich VSL is 25.9 times that of the Poor; but the Poor utilitarian weight is 24.9 times that of the Rich.<sup>10</sup> Thus, while traditional CBA would assign Rich a monetary equivalent for a given small risk reduction (here, a 1 in 100,000 reduction) which is 25.9 times larger than Poor’s monetary equivalent for the same risk reduction, utilitarian weighted CBA would assign Rich an adjusted monetary equivalent which is  $25.9/24.9 = 1.04$  times larger than Poor’s.

Adding isoelastic weights further counteracts the Rich/Poor VSL divergence. Continuing with the scenario of  $\lambda = 2$  and a subsistence level of 1000, note that at a low value (0.5) of

<sup>10</sup> A subsistence level of \$1000 may seem incredibly low, but in fact is in the range of the “extreme poverty” level of \$1.25/day now used by the World Bank and the U.N. On the derivation of this level, see Chen and Ravallion 2010. Empirical evidence on  $\lambda$  is mixed; a value of 2 or even substantially higher is not empirically implausible. See Kaplow 2005.

inequality aversion  $\gamma$ , the Poor/Rich ratio of isoelastic weights is 25.5. Rich's adjusted monetary equivalent is now even closer to Poor's (the ratio is  $25.9/25.5 = 1.02$ ). At larger values of  $\gamma$ , isoelastic weights "overcompensate" for the difference between Rich and Poor VSL: the Poor/Rich ratio of isoelastic weights exceeds the Rich/Poor VSL ratio.

The table also illustrates subtle interactions between  $\lambda$  and  $\gamma$ , which cannot be explored in detail here—except to note that as  $\lambda$  increases and utilitarian weights for the Poor relative to the Rich become larger, the further proportional increase produced by introducing and then increasing inequality aversion is less significant.

The utilitarian and isoelastic weights would also, of course, affect the relative weighting of *income* reductions incurred by Rich or Poor. If someone's income is reduced by amount  $\Delta c$ , traditional CBA assigns the same value ( $\Delta c$ ) to that reduction, regardless of the individual's attributes. Weighted CBA assigns the reduction a value equaling  $\Delta c$  multiplied by the distributional weight. Thus the ratio between the weighted value of a reduction in Poor's income, and the weighted value of the very same reduction in Rich's income, is simply the ratio of distributional weights—as displayed in the fourth row of the table for the utilitarian case, and in subsequent rows for isoelastic weights.

Let us now consider four types of policies, assuming that each group has the same number of members: (I) *Uniform Risk Reduction and Cost Incidence* (both Rich and Poor individuals receive a 1-in-100,000 fatality risk reduction, and incur some reduction in income, the same for both groups); (II) *Uniform Risk Reduction and Redistributive Incidence* (both Rich and Poor individuals receive the risk reduction, but all the costs are borne by the Rich group); (III) *Concentrated Risk Reduction and Cost Incidence* (Poor individuals receive a 1-in-100,000 risk reduction, and pay the costs); and (IV) *Regressive Risk Transfer* (Poor individuals suffer an *increase* in risk of 1-in-100,000, with Rich individuals receiving a risk reduction of the same amount—as would occur with a decision to site a hazardous facility in a geographic location that is closer to where Poor rather than Rich individuals reside).

Table 7 shows how each of these policy choice would be evaluated by traditional CBA, summing monetary equivalents; CBA with population-average rather than differentiated VSL values; CBA with utilitarian weights; and CBA with isoelastic weights. In the case of Uniform Risk Reduction and Cost Incidence, utilitarian CBA is only willing to impose a relatively low uniform income reduction (as compared to traditional or population-average CBA), and isoelastic CBA a yet lower income reduction: a larger reduction would still be net beneficial for the Rich (given their larger VSL values), but too large a net welfare loss for Poor. Conversely, in the case of Uniform Risk Reduction and Redistributive Incidence, utilitarian CBA is willing to impose a larger income reduction (now borne by Rich) than traditional or population-average CBA, and isoelastic CBA yet larger.

**Table 7: The Effect of Distributional Weights on Different Kinds of Risk Policies**

	<b>Uniform Risk Reduction and Cost Incidence</b>  <u>Maximum per capita cost imposed uniformly on Rich and Poor</u>	<b>Uniform Risk Reduction and Redistributive Incidence</b>  <u>Maximum per capita cost imposed on Rich</u>	<b>Concentrated Risk Reduction and Cost Incidence</b>  <u>Maximum per capita cost imposed on Poor</u>	<b>Regressive Risk Transfer</b>  “Yes” if the transfer is assigned a positive sum of monetary equivalents, “No” if it is assigned a negative sum, “Neutral” if assigned a zero sum
CBA w/o weights	\$51.67	\$103.33	\$3.84	<b>Yes</b> $VSL^{Rich}/VSL^{Poor} = 25.9 > 1$
CBA with utilitarian weights	\$7.54	\$194.97	\$3.84	<b>Yes</b> $(VSL^{Rich}/VSL^{Poor}) \div (U-Weight^{Poor}/U-Weight^{Rich}) = 1.04 > 1$
CBA with isoelastic weights, $\gamma = 0.5$	\$7.45	\$197.21	\$3.84	<b>Yes</b> $(VSL^{Rich}/VSL^{Poor}) \div (Iso-Weight^{Poor}/Iso-Weight^{Rich}) = 1.02 > 1$
CBA with isoelastic weights, $\gamma = 1$	\$7.37	\$199.50	\$3.84	<b>No</b> $(VSL^{Rich}/VSL^{Poor}) \div (Iso-Weight^{Poor}/Iso-Weight^{Rich}) = 0.99 < 1$
CBA with isoelastic weights, $\gamma = 2$	\$7.22	\$204.23	\$3.84	<b>No</b> $(VSL^{Rich}/VSL^{Poor}) \div (Iso-Weight^{Poor}/Iso-Weight^{Rich}) = 0.95 < 1$
CBA with population average VSL	\$42.33	\$84.67	\$42.33	<b>Neutral</b> Cost to Poor = \$42.33 $(VSL^{Avg}/100000) = \text{Benefit to Rich}$

Note: These calculations assume the scenario of  $\lambda = 2$  and a subsistence level of 1000. The population-average VSL of \$4.233 million is calculated by assuming that individuals are uniformly distributed from income levels of Poor = \$20,000 in \$10,000 increments to Rich = \$100,000; and that background fatality risk decreases linearly from the Poor level of 0.01 to the Rich Level of 0.005.

In the case of Concentrated Risk Reduction and Cost Incidence, population-average CBA violates the Pareto principle: government may end up imposing an income loss much greater than the Poor are willing to pay for a 1-in-100,000 risk reduction. (This break-even income loss is approximately  $VSL_{Poor}$  times the risk reduction, which in the scenario covered by the table is \$3.84.) Note that traditional CBA avoids this unpleasant result, but so do the weighted versions.

Although the use of population-average CBA conflicts with the Pareto principle, and more generally lacks any theoretical basis, it does have a key intuitive advantage: in a case such as Regressive Risk Transfer, traditional CBA approves the transfer while population-average CBA is neutral. Interestingly, utilitarian CBA also approves the regressive transfer, and isoelastic CBA will also do so for a sufficiently low value of inequality aversion—but as that level becomes higher isoelastic SWF “switches” and favors a progressive risk transfer from Poor to Rich.

## **Heterogeneity of Preferences**

One cluster of worries about distributional weights arises from heterogeneous preferences. Here, we should distinguish between two kinds of heterogeneity: of moral preferences, and of personal preferences.

Heterogeneity of moral preferences is no threat to the view of distributional weighting presented here. Some citizens or officials will oppose CBA as a criterion for assessing governmental policies, while others will endorse CBA. Within the latter group, some will be persuaded by the Kaldor-Hicks defense of CBA, while other will find the SWF framework more attractive. Weighting is a procedure that operationalizes the moral preferences of this last group in a systematic form.

Thus critiques of distributional weights as “value laden” are inapposite. Of course the decision to use weights, and the choice of SWF to be mirrored by weights, are “value laden”—but so is the decision to use CBA at all, or to use the unweighted variant. These decisions, too, involve a whole series of contestable moral judgments. Governmental officials inevitably make such judgments, or work for higher-ups who do.

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. Suffice it to say that the advice welfare economists and moral theorists provide about the specification of weights is addressed to officials with the legal and democratic authority to act on such advice—whoever exactly those officials may be.

Heterogeneity of *personal*, as opposed to *moral*, preferences poses more of a threat to the SWF framework and thus distributional weights. Why is that? The SWF approach, and welfare economics more generally, has always adopted a preference-based view of well-being. Utilities are a measure of preference-satisfaction. More specifically, if one identifies well-being with the satisfaction of *personal* (self-interested) preferences, then utilities measure the satisfaction of

these preferences. But heterogeneity of personal preferences constitutes a real challenge to the construction of utilities that are interpersonally comparable.

With common personal preferences, a interpersonally comparable utility function unique up to a ratio transformation can be constructed (as discussed above) by identifying a vNM utility function representing these common preferences—a function then “zeroed out” by assigning zero to a threshold bundle. However, with heterogeneous personal preferences, there is no *single* preference structure to be represented in this manner.

Using each person’s utility function to quantify *her* bundles is clearly a nonstarter. The difficulty is that these individual functions are not unique. Assume that  $u_{\text{Jim}}(\cdot)$  is a vNM function that represents Jim’s preferences over lotteries, and assigns zero to what Jim regards as a threshold bundle. Any other function  $v_{\text{Jim}}(\cdot)$  which equals  $u_{\text{Jim}}(\cdot)$  multiplied by a positive constant will *also*, then, be a function representing Jim’s preferences over lotteries, and assigning zero to Jim’s threshold bundle. For any assignment of utility numbers to Jim’s attributes, we could equally well multiply those numbers by a positive constant  $a_{\text{Jim}}$ . And this would defeat interpersonal comparability, as illustrated above in Table 2.

Welfare economists *have* developed better proposals for making interpersonal comparisons notwithstanding preference heterogeneity. (For overviews, see Boadway 2012, pp. 199-217; Adler and Fleurbaey, forthcoming). This is an ongoing area of research, which I lack space to review in depth here—except to say it has direct implications for distributional weighting, with each of the various proposals corresponding to a particular methodology for setting weights in the presence of divergent personal preferences.

One possibility is to expand the space of attributes. Heterogeneity with respect to a subset of attributes might reflect a deeper commonality. For example, married persons might have different preferences over health-consumption bundles than single persons, but such divergence is consistent with common preferences over bundles consisting of health, consumption, *and* marital status. Although heterogeneity at the deepest level cannot be ruled out, expanding the space of attributes may at least reduce the degree of heterogeneity.

A second possibility is to pool utility functions. If  $\mathbf{U} = \{u(\cdot), u^*(\cdot), u^{**}(\cdot), \dots\}$  is the set of utility functions each member of which represents the personal preferences of some subset of the population, then we assign an *array* of weights—one for each member of  $\mathbf{U}$ . A policy is decisively better only if better according to all the weights. This has the advantage of being a straightforward generalization of the case of common preferences—the disadvantage of creating zones of indeterminacy where weighting schemes disagree, perhaps large such zones if preference heterogeneity is substantial.

A third proposal is to make “equivalent income” the measure of well-being. The decisionmaker identifies a reference level of prices and reference bundle of non-income attributes,  $\mathbf{p}^{\text{ref}}$  and  $\mathbf{a}^{\text{ref}}$ . Someone’s equivalent income for a given bundle  $(c, \mathbf{p}, \mathbf{a})$ , depends upon

her preferences. It is the amount,  $c^{\text{equiv}}$ , such that she would be indifferent between  $(c^{\text{equiv}}, \mathbf{p}^{\text{ref}}, \mathbf{a}^{\text{ref}})$  and  $(c, \mathbf{p}, \mathbf{a})$ . Recent work has shown how equivalent income can be used to calculate distributional weights. (Fleurbaey, Luchini, Muller, and Schokkaert, 2013).

A fourth proposal builds upon John Harsanyi's concept of "extended preference." (Harsanyi 1986, ch. 4). If  $\mathbf{B}$  is an ordinary bundle, let a hybrid bundle be a combination of an ordinary bundle and a personal preference structure  $R$ . We can imagine the decisionmaker ranking hybrid bundles, and lotteries over hybrid bundles—under the constraint that if someone with personal preferences  $R$  prefers bundle  $\mathbf{B}$  to bundle  $\mathbf{B}^*$ , then the decisionmaker should prefer bundle  $(\mathbf{B}, R)$  to bundle  $(\mathbf{B}^*, R)$ . An extended utility function represents these extended preferences, and can—in principle—be used to assign utility numbers for an SWF, and to generate distributional weights. This approach to weighting, like the equivalent-income approach, will take account both of individuals' attributes and of their personal preferences; and it can avoid the indeterminacy of the pooling approach. (See Adler and Fleurbaey, forthcoming, for a comparison of equivalent incomes and extended preferences).

There is a final wrinkle to the problem of preferences and preference-heterogeneity. Although I have assumed up to this point that utilities should measure *personal* preferences, this is not a point of universal consensus among economists. Some take the position (explicitly or implicitly) that someone is better off if *any* of her preferences are satisfied—not merely her self-interested preferences, but her fairness preferences, altruistic preferences, and for that matter her moral preferences. (Kaplow and Shavell 2002, 18-24).

The SWF framework, and thus distributional weights, are formally consistent with a wide range of views about the kinds of preferences that should be the basis for utility numbers—whether or not such views are *substantively* plausible. The main formal difficulty occurs when such preferences are heterogeneous. For example, imagine that utilities are supposed to represent the combination of individuals' personal and fairness preferences. (Johnasson-Stenman and Konow, 2008). An individual's "attributes" should therefore describe not only her own consumption, health, etc., but how she fares relative to others. If individuals have common personal-plus-fairness preferences over bundles of these sorts of attributes, then a vNM representation of such preferences can function as the input to an SWF, and as the basis for distributional weights. (Someone's  $MU_i$  weight, specifically, would measure the extent to which a small increase in his consumption enhances his self-interest or his sense of fair treatment.)

It is regularly suggested that distributional goals should be handled by the tax system, with non-tax governmental bodies (such as regulatory agencies or bodies that provide public goods) focused solely on “efficiency.” More precisely, it might be claimed that the case for distributional weights is undermined by the existence of a governmental system for taxing income, consumption, or wealth, as well as paying income subsidies (“the tax system”). In considering this claim, we should distinguish three quite distinct lines of argument.

(1) If taxes-and-transfers are optimal, distributional weights are uniform

Assume that the decisionmaker with legal authority over some non-tax body has moral preferences reflected in a utilitarian or isoelastic SWF, and that the legislature (which controls the tax system) has the very same moral preferences. Assume, further, that the tax system is a “lump sum” system, meaning: (a) There is zero administrative cost to collecting taxes or paying subsidies. (b) The system has no adverse incentive effects: the level of each individual’s labor supply and other non-consumption attributes (such as health), and the total produced stock of each type of marketed good and service, is independent of the level of taxes and subsidies. (c) Each individual’s attributes are observable to the officials who administer the tax system.

With all these assumptions in place, utilitarian or isoelastic distributional weights *will* be identical for all individuals. (See Dreze and Stern 1987, pp. 935-37; Salanie 2003, pp. 79-83; Appendix.) In the case where the legislature and decisionmaker share utilitarian moral preferences, the legislature will have enacted taxes-and-transfers that equalize individuals  $MU_i$  values. And in the case where they have isoelastic moral preferences (with the same value of  $\gamma$ ), the legislature will have enacted taxes-and-transfers that equalize individuals’  $MU_i \times MMVU_i$  values. In either case, the decisionmaker can maximize her preferred SWF by undertaking CBA without weights.

However, the assumptions can readily fail. First, the decisionmaker may have different moral preferences from the legislature’s, or some members of the legislature. For example, a President elected by one political party may have more strongly egalitarian preferences than the competing party that controls the legislature. He embraces a isoelastic SWF, while they are utilitarians. He is utilitarian, while they are libertarians. If so, and even if the tax system is a “lump sum” system, the status quo distribution of income could well be such that distributional weights *according to the SWF that the President endorses* are unequal. Thus the President (given his moral preferences) would have good moral reason to direct non-tax bodies within his control to use CBA with distributional weights; and he might well also believe that doing so is democratically and legally legitimate.

It is worth noting, here, that a diversity of moral preferences (on the part of diverse governmental decisionmakers) is reflected in the actual institutional history of CBA in the U.S. Since 1981, a Presidential order has directed Executive Branch agencies to use CBA as a general tool for evaluating regulation; but proponents of CBA have never succeeded in persuading

Congress to embody this directive in a statute. (Wiener 2006, pp. 461-66.) Thus it is hardly science fiction to suppose that a President might prefer to move CBA, now in place, in a more egalitarian direction, but be unable to secure tax-code changes reflecting this preference.

Second, even if the legislature and decisionmaker share a moral commitment to a utilitarian SWF or to an isoelastic SWF with a particular value of  $\gamma$ , the absence of a lump-sum tax-and-transfer system may mean that distributional weights in the status quo are *not* equal by the lights of the shared SWF. The most systematic scholarly thinking about non-lump-sum taxation has occurred in the field of optimal tax theory, where it is traditionally assumed that individuals' labor incomes are observable by government but that their underlying abilities (yielding higher wages for more able individuals) are not; that higher income taxes will disincentive labor at the margin, since individuals have a preference for leisure as well as for the consumption of marketed goods; and that the tax code is set to maximize an SWF, e.g., a utilitarian or isoelastic one. (See, e.g., Boadway 2012; Kaplow 2008; Salanie 2003; Tulomala 1990.) It is *not* generally true in this context that the optimal taxes (in light of this SWF) will yield a pattern of income and non-income attributes such that distributional weights (in light of that same SWF) are equal. (Fleurbaey, Luchini, Muller, and Schokkaert 2013; Layard 1980). The optimal-tax setup typically ignores administrative costs; adding these to the tax system may also help to produce unequal distributional weights at the optimum. (See Quiggin 1995.)

Finally, a scheme of distributional weights is a *schedule*, assigning different individuals (or groups) weights as a function of their status quo attributes. If the tax system *has* operated to produce a pattern of attributes such that individuals have uniform  $MU_i$  or  $MU_i \times MMVU_i$  values, the schedule will assign uniform weights, and will reduce to unweighted CBA; otherwise, weights will be unequal. The possibility that weights will be identical under certain conditions is not, itself, an argument against such a schedule.

(2) Whether or not distributional weights are uniform, the decisionmaker can produce a Pareto-superior result by a combination of unweighted CBA and tax changes

A second and more powerful argument against the use of distributionally weighted CBA derives from Hylland and Zeckhauser (1979) and has been developed very fully in recent years by Kaplow (see, e.g., Kaplow 1996, 2008). The argument can be articulated as follows. Consider a non-tax decisionmaker who endorses some Paretian SWF (be it the isoelastic SWF, the utilitarian SWF, or any other SWF that respects the Pareto principle). Assume that some policy  $P^*$  is chosen by traditional CBA without weights, as opposed to an alternative policy  $P$ . Regardless of the pattern of status quo attributes—regardless whether this pattern is such as to produce uniform or nonuniform distributional weights in light of the SWF—it will be Pareto-superior for the decisionmaker to implement policy  $P^*$  together with matching changes to the tax code, as opposed to enacting  $P$ . If  $P^*$  passes a traditional CBA test relative to  $P$ , then those who gain from  $P^*$  have enough to compensate those who lose. Moreover, Kaplow shows how these gains can be redistributed from the winners to the losers to yield a Pareto-superior result *even*

*with a non-lump-sum tax system.* (Specifically, he works within the set-up from optimal tax theory, which takes account of the unobservability of individuals' underlying abilities, and of the labor-disincentive effects of an income tax.) Now note, finally, that if  $P^*$  together with a change to the tax code is Pareto-superior to  $P$ , any Pareto-respecting SWF will favor that package as opposed to the enactment of  $P$ .

Kaplow's analysis relies on a key technical assumption of labor separability, which allows him to construct tax changes that do not change labor supply. This assumption has no clear empirical basis, and a different assumption is adopted by much scholarship in environmental economics. (Johansson-Stenman 2005; Fullerton 2009, p. xxi). Moreover, the analysis does not consider the administrative costs of changes to the tax system, or—once again—the possibility that a non-tax decisionmaker who favors some SWF may not control the tax system. Imagine that a decisionmaker has the power to choose between  $P$  and  $P^*$ . If the tax code does not change, her morally preferred SWF favors  $P$  over  $P^*$ . Although  $P^*$  bundled with a change to the tax code is Pareto-superior to  $P$ , the decisionmaker does not believe the legislature will make this tax change. She is therefore morally justified in picking  $P$ .

Given these limitations, Kaplow's analysis does not demonstrate that a decisionmaker favoring some SWF should refrain from putting in place or following a matching schedule of distributional weights. Rather, the analysis has a less dramatic—but still very important—implication. It suggests that the “policies” being considered—using the schedule—should include not merely non-tax policies, but also such policies combined with changes to the tax code *where the decisionmaker believes these changes to be politically feasible*. In some cases,  $P$  will be favored by distributionally weighted CBA (mirroring the SWF) over  $P^*$ ; but  $P^*$  *together with a politically feasible change to the tax system* will be favored by distributionally weighted CBA over  $P$ , even given the actual administrative costs of this change, and the predicted incentive effects of the change on labor supply (zero or not). If  $P^*$ -plus-the-feasible-tax-change is Pareto superior to  $P$ , distributionally weighted CBA will prefer the first option to the second; and it might do so even absent Pareto superiority.

In short, decisionmakers favoring an SWF mirrored by distributional weights should consider an expanded menu of policy options—to include tax system changes, if feasible, as well as non-tax policies—but the existence of the tax system does not show that they should abandon weighting entirely.

### (3) The adverse incentive effects of distributional weights

Distributional weights might have adverse incentive effects. A lower weight for those with higher income might disincentivize labor or other income-generating activities: by earning more income, an individual gives her interests less weight in the policy process, and thus increases the chance of policies that benefit others at her expense. Similarly, a lower weight for those in better health might discourage healthy lifestyles.

The incentive effects of income tax rates have been intensively studied; by contrast, the incentive effects of distributional weights are little discussed, even within scholarship on weighting. It is possible that heterogeneity in weights would be less salient to the public than heterogeneity in tax rates, so that incentive issues would be less significant in the first case. (Cf. Jolls 1998.) The topic needs more scholarly examination.

In any event, the lesson here is that decisionmakers favoring some SWF must consider incentive effects, and may need to modulate the weights in light of such effects. But this does not mean that optimal weights are *equal*. (Consider that incentive arguments may argue for a less progressive schedule of marginal tax rates, but have certainly not demonstrated that marginal rates should be *equal*.)

## Conclusion

This Article has shown how to “put structure” on the problem of distributional weights, and has addressed some recurrent objections. The decision to use weights to mimic a particular SWF *does* require an ethical/moral judgment, but so does the decision to use unweighted CBA—a procedure that is intensely controversial outside the community of economists. (Ackerman and Heinzerling 2004) The specification of utilitarian weights is quite straightforward; these simply reflect the marginal utility of consumption/income. The specification of “isoelastic” weights is somewhat more complicated—requiring a further ethical judgment regarding the appropriate degree of inequality aversion  $\gamma$ , and the specification of a “zero point” (subsistence level) so that utility will be measurable on a ratio scale. Still, some will feel that the utilitarian SWF is insufficiently sensitive to distribution, and that this additional complexity is justified.

The use of weights will dampen income elasticity in money equivalents for environmental goods (such as fatality risk reduction), without producing a conflict with the Pareto principle. The absence of a lump-sum tax system, and the economic and political costs of tax reform, undercut the objection that distributional concerns should be wholly relegated to the tax code. Perhaps the most powerful objection to weights is that they presume the interpersonal comparability of well-being, which breaks down with heterogeneous preferences. However, recent years have seen theoretical advances on the problem of interpersonal comparisons. More work in this area is critical, and will further refine our understanding of how to set weights

## References

Ackerman, Frank, and Lisa Heinzerling. 2004. *Priceless: On Knowing the Price of Everything and the Value of Nothing*. New York: New Press.

Adler, Matthew D. 2012. *Well-Being and Fair Distribution: Beyond Cost-Benefit Analysis*. Oxford: Oxford University Press.

Adler, Matthew D., and Marc Fleurbaey. Forthcoming. *Oxford Handbook of Well-Being and Public Policy*. Oxford: Oxford University Press.

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

Boadway, Robin. 2012. *From Optimal Tax Theory to Tax Policy*. Cambridge, MA: MIT Press.

Bergson, Abram. 1938. "A Reformulation of Certain Aspects of Welfare Economics." *The Quarterly Journal of Economics* 52: 310-34.

\_\_\_\_\_. 1948. "Socialist Economics." In *A Survey of Contemporary Economics*, ed. Howard S. Ellis, vol. 1, 412-48. Homewood: Richard D. Irwin.

\_\_\_\_\_. 1954. "On the Concept of Social Welfare." *The Quarterly Journal of Economics* 68: 233-52.

Bossert, Walter, and John A. Weymark. 2004. "Utility in Social Choice." In *Handbook of Utility Theory*, ed. Salvador Barbera, Peter J. Hammond, and Christian Seidl, vol. 2 (*Extensions*), 1099-1177. Boston: Kluwer Academic.

Brent, Robert J. 1984. "Use of Distributional Weights in Cost-Benefit Analysis: A Survey of Schools." *Public Finance Quarterly*. 12: 213-30.

Chen, Shaohua, and Martin Ravallion. 2010. "The Developing World is Poorer than We Thought, But No Less Successful in the Fight Against Poverty." *The Quarterly Journal of Economics* 125: 1577-1625.

Chipman, John S., and James C. Moore. 1978. "The New Welfare Economics 1939-74." *International Economic Review*. 19: 547-84.

Cowell, Frank A., and Karen Gardiner. 1999. "Welfare Weights." STICERD, London School of Economics. [http://darp.lse.ac.uk/papersDB/Cowell-Gardiner\\_\(OFT\).pdf](http://darp.lse.ac.uk/papersDB/Cowell-Gardiner_(OFT).pdf).

Creedy, John. 2006. "Evaluating Policy: Welfare Weights and Value Judgments." University of Melbourne Dept. of Economics Working Paper No. 971. <http://ideas.repec.org/p/mlb/wpaper/971.html>.

- Cropper, Maureen, James K. Hammitt, and Lisa A. Robinson. 2011. "Valuing Mortality Risk Reductions: Progress and Challenges." *Annual Review of Resource and Environmental Economics*. 3: 313-36.
- Dasgupta, Ajit K., and D.W. Pearce. 1972. *Cost-Benefit Analysis: Theory and Practice*. New York: Harper and Row.
- Dasgupta, Partha, Amartya Sen, and Stephen Marglin. 1972. *Guidelines for Project Evaluation*. Vienna: United Nations Industrial Development Organization.
- Dasgupta, Partha. 2008. "Discounting Climate Change." *Journal of Risk and Uncertainty* 37: 141-69.
- Dreze, Jean, and Nicholas Stern. 1987. "The Theory of Cost-Benefit Analysis." In *Handbook of Public Economics*, ed. Alan J. Auerbach and Martin Feldstein, vol. 2, 909-89. Amsterdam: North-Holland.
- Dreze, Jean. 1998. "Distribution Matters in Cost-Benefit Analysis: Comment on K.A. Brekke." *Journal of Public Economics*. 70: 485-88.
- Edwards, C.H., Jr. 1994. *Advanced Calculus of Several Variables*. New York: Dover Publications. Corrected republication of 1973 edition published by Academic Press.
- Eeckhoudt, Louis R., and James K. Hammitt. 2004. "Does Risk Aversion increase the Value of Mortality Risk?" *Journal of Environmental Economics and Management*. 47: 13-29.
- Evans, Mary F., and V. Kerry Smith. 2010. "Measuring How Risk Tradeoffs Adjust with Income." *Journal of Risk and Uncertainty*. 40: 33-55.
- Evans, William N., and W. Kip Viscusi. 1991. "Estimation of State-Dependent Utility Functions using Survey Data." *The Review of Economics and Statistics*. 73: 94-104.
- Finkelstein, Amy, Erzo F.P. Luttmer, and Matthew J. Notowidigdo. 2008. "What Good is Wealth Without Health? The Effect of Health on the Marginal Utility of Consumption." NBER Working Paper No. 14089. <http://www.nber.org/papers/w14089.pdf>
- Fleurbaey, Marc. 2010. "Assessing Risky Social Situations." *Journal of Political Economy*. 118: 649-80.
- Fleurbaey, Marc, Stephane Luchini, Christophe Muller, and Erik Schokkaert. 2013. "Equivalent Income and Fair Evaluation of Health Care." *Health Economics*. 22: 711-29.
- Freeman, A. Myrick. 2003. *The Measurement of Environmental and Resource Values: Theory and Methods*. 2<sup>nd</sup> ed. Washington, D.C.: Resources for the Future.

- Fullerton, Don, ed. 2009. *Distributional Effects of Environmental and Energy Policy*. Farnham: Ashgate.
- Gollier, Christian. 2001. *The Economics of Risk and Time*. Cambridge, MA: MIT Press.
- Gorman, W.M. 1955. "The Intransitivity of Certain Criteria used in Welfare Economics." *Oxford Economic Papers*. 7: 25-35.
- Harberger, Arnold. 1978. "On the Use of Distributional Weights in Social Cost-Benefit Analysis." *Journal of Political Economy*. 86: S87-S120.
- Harsanyi, John. 1977. *Rational Behavior and Bargaining Equilibrium in Games and Social Situations*. Cambridge: Cambridge University Press.
- HM Treasury. 2003. *The Green Book: Appraisal and Evaluation in Central Government*. Norwich: The Stationery Office.
- Hylland, Aanund, and Richard Zeckhauser. 1979. "Distributional Objectives should Affect Taxes but not Program Choice or Design." *Scandinavian Journal of Economics*. 81: 264-84.
- Johansson-Stenman, Olof. 2005. "Distributional Weights in Cost-Benefit Analysis: Should We Forget about Them?" *Land Economics*. 81: 337-52.
- Johansson-Stenman, Olof. 2000. "On the Value of Life in Rich and Poor Countries and Distributional Weights beyond Utilitarianism." *Environmental and Resource Economics*. 17: 299-310.
- Johansson-Stenman, Olof, and James Konow. 2010. "Fair Air: Distributive Justice and Environmental Economics." *Environmental and Resource Economics*. 46: 147-66.
- Jolls, Christine. 1998. "Behavioral Economics Analysis of Redistributive Legal Rules." *Vanderbilt Law Review*. 51: 1653-77.
- Kaplow, Louis. 2008. *The Theory of Taxation and Public Economics*. Princeton: Princeton University Press.
- Kaplow, Louis. 1996. "The Optimal Supply of Public Goods and the Distortionary Cost of Taxation." *National Tax Journal*. 49: 513-33.
- Kaplow, Louis. 2005. "The Value of a Statistical Life and the Coefficient of Relative Risk Aversion." *Journal of Risk and Uncertainty*. 31: 23-34.
- Kaplow, Louis, and Steven Shavell. 2002. *Fairness versus Welfare*. Cambridge, MA: Harvard University Press.

- Kenney, Ralph, and Howard Raiffa. 1993. *Decisions with Multiple Objectives: Preferences and Value Tradeoffs*. Cambridge: Cambridge University Press.
- Lambert, Peter J., Daniel L. Millimet and Daniel Slottje. 2003. "Inequality Aversion and the Natural Rate of Subjective Inequality." *Journal of Public Economics*. 87: 1061-90.
- Layard, Richard, 1980. "On the Use of Distributional Weights in Social Cost-Benefit Analysis." *Journal of Political Economy*. 88: 1041-47.
- Liu, Liqun. 2006. "Combining Distributional Weights and the Marginal Cost of Funds." *Public Finance Review*. 34: 60-78.
- Kreps, David M., 1988. *Notes on the Theory of Choice*. Boulder: Westview Press.
- Little, I.M.D., and J.A. Mirrlees. 1974. *Project Appraisal and Planning for Developing Countries*. New York: Basic Books.
- Little, I.M.D., and J.A. Mirrlees. 1994. "The Costs and Benefits of Analysis: Project Appraisal and Planning Twenty Years On." In *Cost-Benefit Analysis*, ed. Richard Layard and Stephen Glaister, 199-231. Cambridge: Cambridge University Press.
- Mas-Colell, Andreu, Michael Whinston and Jerry R. Green. 1995. *Microeconomic Theory*. New York: Oxford University Press.
- Meade, J.E. 1955. *Trade and Welfare: Mathematical Supplement*. London: Oxford University Press.
- Nordhaus, William. 2008. *A Question of Balance: Weighing the Options on Global Warming Policies*. New Haven: Yale University Press.
- Office of Management and Budget. 2003. *Circular A-4: Regulatory Analysis*. Available at [http://www.whitehouse.gov/omb/inforeg\\_default/](http://www.whitehouse.gov/omb/inforeg_default/)
- Parry, Ian, et al. 2006. "The Incidence of Pollution Control Policies." In *International Yearbook of Environmental and Resource Economics 2006/07*, ed. Henk Folmer and Thomas Tietenberg. Cheltenham: Edward Elgar
- Pauwels, W. 1978. "The Possible Perverse Behavior of the Compensating Variation as a Welfare Ranking." *Zeitschrift für Nationalökonomie/Journal of Economics*. 38: 369-78.
- Quiggin, John. 1995. "The Suboptimality of Efficiency." *Economics Letters*. 47: 389-92.
- Ray, Anandarup. 1984. *Cost-Benefit Analysis: Issues and Methodologies*. Baltimore: Johns Hopkins University Press.

- Rey, Beatrice, and Jean-Charles Rochet. 2004. "Health and Wealth: How do They Affect Individual Preferences?" *The Geneva Papers on Risk and Insurance Theory*. 29: 43-54.
- Salanie, Bernard. 2002. *The Economics of Taxation*. Cambridge, MA: MIT Press.
- Samuelson, Paul A. 1947. *Foundations of Economic Analysis*. Cambridge, MA: Harvard University Press.
- Savage, Leonard J. 1972. *The Foundations of Statistics*. 2<sup>nd</sup> revised edition. New York: Dover.
- Sen, Amartya. 1970. *Collective Choice and Social Welfare*. San Francisco: Holden-Day.
- Sen, Amartya. 1979. "The Welfare Basis of Real Income Comparisons: A Survey." *Journal of Economic Literature*. 17: 1-45.
- Scitovsky, T. 1941. "A Note on Welfare Propositions in Economics." *The Review of Economic Studies*. 9: 77-88.
- Squire, Lyn, and Herman G. van der Tak. 1975. *Economic Analysis of Projects*. Baltimore: Johns Hopkins University Press.
- Stern, Nicholas. 2007. *The Economics of Climate Change: The Stern Review*. Cambridge: Cambridge University Press.
- Sunstein, Cass. 2004. "Valuing Life: A Plea for Disaggregation." *Duke Law Journal*. 54: 385-445.
- Thomas, George B., Jr., and Ross L. Finney. 1998. *Calculus and Analytic Geometry*. 9<sup>th</sup> ed. Reading: Addison-Wesley.
- Tuomala, Matti. 1990. *Optimal Income Tax and Redistribution*. Oxford: Clarendon Press.
- Viscusi, W. Kip. 2010. "The Heterogeneity of the Value of Statistical Life: Introduction and Overview." *Journal of Risk and Uncertainty*. 40: 1-13.
- Wiener, Jonathan B. 2006. "Better Regulation in Europe." *Current Legal Problems*. 59: 447-518.
- Weisbrod, Burton A. 1968. "Income Redistribution Effects and Benefit-Cost Analysis." In *Problems in Public Expenditure Analysis*, ed. Samuel B. Chase, 177-209. Washington, D.C.: The Brookings Institution.
- Yitzhaki, Shlomo. 2003. "Cost-Benefit Analysis and the Distributional Consequences of Government Projects." *National Tax Journal*. 56: 319-36.

## Appendix for Adler, “Cost-Benefit Analysis and Distributional Weights: An Overview”

### I. Utilitarian Distributional Weights under Certainty

#### A. Marginal weights with equivalent variations

Let  $\mathbf{B}_i^s = (c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)$  be individual  $i$ 's bundle of attributes in the status quo  $s$ —consumption, prices, and nonconsumption attributes—and  $\mathbf{B}_i^x = (c_i^x, \mathbf{p}^x, \mathbf{a}_i^x)$  individual  $i$ 's bundle in some alternative outcome  $x$ . Assume common personal preferences. Let  $v(\cdot)$  be a vNM utility function representing those preferences, with respect to bundles of marketed goods and nonconsumption attributes; and let  $u(\cdot) = u(c, \mathbf{p}, \mathbf{a})$  be the corresponding indirect utility function, i.e., the maximum value of  $v(\cdot)$  achievable with  $(c, \mathbf{p}, \mathbf{a})$ . Let  $\Delta c_i^x$  be individual  $i$ 's equivalent variation for  $x$ , relative to the status quo, such that:  $u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s) = u(c_i^x, \mathbf{p}^x, \mathbf{a}_i^x)$ .

Let  $w(\cdot)$  here denote the utilitarian social welfare function.  $w(x) = \sum_{i=1}^N u(\mathbf{B}_i^x)$ .<sup>11</sup> This SWF orders outcomes using the rule:  $x$  at least as good as  $y$  iff  $w(x) \geq w(y)$ . That ordering, in turn, is identical to using the rule:  $x$  at least as good as  $y$  iff  $w(x) - w(s) \geq w(y) - w(s)$ . Note that  $w(x) - w(s) = \sum_{i=1}^N [u(c_i^x, \mathbf{p}^x, \mathbf{a}_i^x) - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)] = \sum_{i=1}^N [u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s) - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)]$ . Assume that  $u(\cdot)$  is continuously differentiable. Then this last sum is well approximated, for equivalent variations sufficiently small, with the standard total-differential approximation:

$$\sum_{i=1}^N \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} \Delta c_i^x, \text{ with } \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} = \text{MU}_i \text{ denoting the partial derivative of } u(\cdot) \text{ with}$$

respect to  $c$  at the value  $(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)$ . (The total-differential approximation to a continuously differentiable real-valued function is discussed in any good calculus textbook. See, e.g., Thomas and Finney 1998, pp. 933-44; Edwards 1994, pp. 63-76.)

#### B. Compensating variations

CBA using compensating variations relative to the status quo can violate the Pareto principle. See Freeman 2003, pp. 61-62; Pauwels 1978. This problem carries over to CBA with compensating variations and distributional weights. The conflict with Pareto arises because variation in prices or nonconsumption attributes across the outcomes being compared can change the marginal utility of consumption.

For a simple example, assume that there are three outcomes: the status quo,  $x$ , and  $y$ . Prices are the same in all outcomes. Individuals have a consumption amount and a binary health attribute, sick or healthy.  $u(\cdot)$  takes the form:  $u(c, \text{healthy}) = c$ ;  $u(c, \text{sick}) = 2c/5$ . Thus utility is linear in consumption in both states, but the marginal utility of consumption is lower when sick,

<sup>11</sup> To write the SWF as a function of indirect utility  $u(\cdot)$  depends upon an assumption of individual rationality (that each individual in fact maximizes her direct utility given her expenditure budget and the price vector) and market equilibrium (that it is possible for all individuals to do without prices changing). These are, of course, pervasive assumptions in neoclassical economics.

and an individual would be indifferent, for example, between the bundle (\$40,000, healthy) and (\$100,000, sick).

Assume that in the status quo one individual, Dylan, has consumption \$149,900 and is sick. In outcome  $x$  he has consumption \$60,000 and is healthy. In outcome  $y$  he has consumption \$149,970 and is sick. Note now that Dylan prefers  $x$  to  $y$  (given the utility function representing common preferences, including his); but that his compensating variation for  $x$  is \$40, while his compensating variation for  $y$  is larger, \$70. Assume no one else is affected. Then CBA summing compensating variations prefers  $y$  to  $x$ , in violation of Pareto superiority.

Nothing changes if we introduce distributional weighting based on status quo attributes, since the compensating variations in  $x$  and  $y$  for Dylan will be multiplied by the same weight (in this case,  $MU_i$  for Dylan is  $2/5$ , because Dylan is sick in the status quo).

The problem is “inframarginal” with respect to the underlying attributes, since we have jumps in Dylan’s consumption and health state between  $x$  and the other two outcomes. However, note that the compensating variations themselves are small in the practical sense, and we can construct examples to make them arbitrarily small. For any arbitrarily small positive amount  $CV$ , let Dylan’s consumption in the status quo be  $\$150,000 - CV - \varepsilon$  (with  $\varepsilon$  a positive number less than  $3/2 CV$ ), and let his consumption in  $y$  be  $\$150,000 - \varepsilon$ . His consumption in  $x$  remains \$60,000. Then it can be seen that Dylan’s compensating variation for  $y$  ( $CV$ ) is larger than his compensating variation for  $x$  ( $2CV/5 + 2\varepsilon/5$ ), and yet he is better off in  $x$ .

By contrast (returning to the original example), note that Dylan’s equivalent variation for  $x$  is \$100, while his equivalent variation for  $y$  is \$70. Summing equivalent variations (with or without distributional weights) satisfies the Pareto principle.

Weighting compensating variations by outcome-specific distributional weights would cure the violation of the Pareto principle, but is much more complicated than assigning each person a single weight dependent just on her status quo attributes, and multiplying her equivalent variation for any other outcome by that weight.

### C. Linear transformations of $u(\cdot)$

If  $u(\cdot)$  is a vNM function representing common personal preferences over bundle lotteries, then so is  $u^*(\cdot) = au(\cdot) + b$ ,  $a$  positive. Note that  $MU_i$  according to  $u^*(\cdot)$  is a positive multiple (by  $a$ ) of  $MU_i$  according to  $u(\cdot)$ .  $\frac{u^*(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$   $a \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$  But of course

$$\sum_{i=1}^N \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} c_i^x \quad \sum_{i=1}^N \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} c_i^y \text{ iff}$$

$$\prod_{i=1}^N a \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} c_i^x \quad \prod_{i=1}^N a \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} c_i^y .$$

Thus using either set of weights orders outcomes exactly the same way.

#### D. Simplifications of $MU_i$ : Invariance and CRRA

For simplicity we are assuming, throughout, that individuals face the same prices in the status quo. We abbreviate  $u(\cdot)$ , given status quo prices, as  $u(c, \mathbf{a})$  and assume, throughout the analysis to follow, that prices are fixed at the status quo level  $\mathbf{p}^s$ .

Assume that the common personal preferences are such that gambles over consumption, holding fixed non-consumption attributes, are invariant to the level of non-consumption attributes. Let  $h(\cdot)$  be a vNM function of consumption that represents these invariant preferences. Let  $u(\cdot, \mathbf{a})$  denote the utility values assigned by  $u(\cdot)$  to bundles with non-consumption attributes fixed at level  $\mathbf{a}$ . By vNM theory,  $u(\cdot, \mathbf{a})$  must be a linear transformation of  $h(\cdot)$ . That is, there must exist a positive constant  $m(\mathbf{a})$  and a constant  $k(\mathbf{a})$  such that  $u(c, \mathbf{a}) = m(\mathbf{a}) h(c) + k(\mathbf{a})$ . (See Keeney and Raiffa, ch. 5). Thus  $\frac{u(c_i^s, \mathbf{a}_i^s)}{c} = h(c_i^s) m(\mathbf{a}_i^s)$

Assume now that  $h(\cdot)$  takes the CRRA form.  $h(c) = (1-\lambda)^{-1} c^{1-\lambda}$ , or  $\ln c$  in the case of  $\lambda = 1$ . Then  $h'(c) = c^{-\lambda}$ , and  $\frac{u(c_i^s, \mathbf{a}_i^s)}{c} = (c_i^s)^{-\lambda} m(\mathbf{a}_i^s)$ .

Note finally that  $m(\cdot)$  will be unity if  $u(\cdot)$  takes the additively separable form  $u(c, \mathbf{a}) = h(c) + k(\mathbf{a})$ . Thus with this additively separable form,  $MU_i$  is *just* a function of individual  $i$ 's status quo consumption.

To be sure,  $u(\cdot)$  is only unique up to a linear transformation; its ranking of bundle lotteries is equally well represented by  $u^*(\cdot) = au(\cdot) + b$ . If  $u(c, \mathbf{a}) = m(\mathbf{a}) h(c) + k(\mathbf{a})$ , then  $u^*(c, \mathbf{a}) = am(\mathbf{a}) h(c) + ak(\mathbf{a}) + b$ . Thus  $MU_i$  according to  $u^*(\cdot)$  also takes the simple form that follows

from invariance:  $\frac{u^*(c_i^s, \mathbf{a}_i^s)}{c} = h(c_i^s) m^*(\mathbf{a}_i^s)$ , with  $m^*(\cdot) = am(\cdot)$ .

#### E. Inframarginal weights

Regardless of the magnitude of equivalent variations, the utilitarian ordering of outcomes can be precisely (not approximately) mirrored using inframarginal weights. Assume that, for all outcomes other than the status quo itself, equivalent variations are nonzero. Let  $\theta_i$  be a weighting factor for individual  $i$  which is a function both of  $i$ 's status quo attributes *and* his equivalent variation. (Note that, by contrast,  $MU_i$  is a function just of  $i$ 's status quo attributes).

Specifically, let  $\theta_i = \theta(\mathbf{B}_i^s, \Delta c_i^x) = [u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s) - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)] / \Delta c_i^x$ . Then obviously  $w(x) - w(s) = \sum_{i=1}^N c_i^x [u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s) - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)] / c_i^x = \sum_{i=1}^N c_i^x \pi(\mathbf{B}_i^s, c_i^x)$

A different and equally correct treatment of the inframarginal case, sometimes seen in the literature, is as follows. Consider  $MU_i$  as a function of individual  $i$ 's consumption  $c_i$ , with prices and  $i$ 's nonconsumption attributes fixed at the status quo level.  $MU_i(c_i) = \frac{u(c_i, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$ . Then

$$u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s) - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) = \int_{c_i^s}^{c_i^s + \Delta c_i^x} MU_i(c_i) dc_i. \quad \text{The distributional weight for individual } i$$

is this integral divided by  $\Delta c_i^x$ . Note also that, by the mean value theorem, this weight is in turn equal to  $MU_i(c_i^*)$  at some value  $c_i^*$  in between  $c_i^s$  and  $c_i^s + \Delta c_i^x$ .

## II. Isoelastic Distributional Weights under Certainty

### A. Marginal weights for equivalent variations

Let  $e(\cdot)$  denote the isoelastic/Atkinson social welfare function. With  $\gamma > 0$ ,  $e(x) = (1 + \gamma)^{-1} \sum_{i=1}^N u(\mathbf{B}_i^x)^{1+\gamma}$ , or  $\sum_{i=1}^N \ln u(\mathbf{B}_i^x)$  if  $\gamma = 1$ . On the properties of this SWF, see generally Adler (2012, ch. 5).

The construction of distributional weights directly parallels the construction above of utilitarian distributional weights. With  $s$  the status quo, the ranking of outcomes by  $e(\cdot)$  is identical to their ranking by  $e(\cdot) - e(s)$ . But  $e(x) - e(s) =$

$$\sum_{i=1}^N (1 + \gamma)^{-1} [u(c_i^s + \Delta c_i^x, \mathbf{p}^s, \mathbf{a}_i^s)^{1+\gamma} - u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)^{1+\gamma}]. \quad \text{Using the total differential approximation}$$

(and the chain rule), this sum is approximated by  $\sum_{i=1}^N u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} \Delta c_i^x$ , with  $MU_i$

as before equaling  $\frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$  and the marginal moral value of utility (MMVU <sub>$i$</sub> ) equaling  $u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)$ . This is the change in the value of  $e(\cdot)$  with a small change in  $i$ 's utility. (To see this most directly, write  $e(\cdot)$  as  $e(u_1, \dots, u_N) = (1 + \gamma)^{-1} \sum_{i=1}^N u_i^{1+\gamma}$ . Then  $\frac{\partial e}{\partial u_i} = u_i^\gamma$ .)

In order for the isoelastic SWF to be not only well-defined, but *Paretian* (increasing in each individual's utility) and *equity-regarding* (the sum of a strictly concave function of individual utilities, so that greater weight is given to welfare changes at lower utility levels), all utility values must be nonnegative, and strictly positive with  $\gamma \geq 1$ . Utility values can be zero or

positive for  $\gamma < 1$ , e.g., where the Atkinson function is the sum of the square root of utilities ( $\gamma = 1/2$ ).

But note that even with  $\gamma < 1$ , *distributional weights* for individual  $i$  are undefined if  $u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) = u_i = 0$ . Thus isoelastic *distributional weights* require that all individuals in the status quo have strictly positive utilities, for all values of  $\gamma$ .

While utility levels of zero lead to undefined isoelastic distributional weights, negative utility values lead to undefined distributional weights *or* well-defined weights that are “badly behaved.” Consider the three cases of  $\gamma = 1/2$ ,  $\gamma = 1/3$ , and  $\gamma = 2$ . In the first case, the MMVU $_i$  term is  $(u_i)^{-1/2} = 1/(u_i)^{1/2}$ , which is undefined for negative values of  $u_i$ . In the second case, that term is  $1/(u_i)^{1/3}$ , which is well-defined but is, perversely, *negative*—with the result that utility improvements to those at negative levels in the status quo would be seen by weighted CBA as social *losses*! In the third case, it is  $1/(u_i)^2$ , which is well-defined and positive. However, these weights are increasing with negative inputs. If  $u_j < u_i < 0$ ,  $1/(u_j)^2 < 1/(u_i)^2$ . Thus weighted CBA gives *less* weight to well-being changes affecting  $j$  than  $i$ , even though  $j$  is worse off than  $i$ .

These difficulties in isoelastic weights with negative utilities just reflect the fact that the isoelastic SWF with negative utilities in its domain is not well-defined *and* Paretian *and* equity-regarding.

## B. Ratio transformations

The isoelastic SWF  $e(\cdot)$  is not invariant to linear transformations of utility. Note that, if  $u^*(\cdot) = au(\cdot) + b$ , with  $a$  positive and  $b$  possibly nonzero,  $e(\cdot)$  using  $u^*(\cdot)$  does *not* necessarily order outcomes the same way as  $e(\cdot)$  using  $u(\cdot)$ . Indeed, it can be shown that the only Paretian, anonymous, separable SWFs that are invariant to linear transformations of utility are the leximin, leximax, and utilitarian SWFs, as well as some other SWFs that are within the class of weakly utilitarian SWFs (those that agree with the utilitarian SWF in ranking vectors with different total amounts of utility). See Bossert and Weymark 2004, pp. 1157-58.

However, the isoelastic SWF *is* invariant to ratio transformations of utility. (It is the only continuous prioritarian SWF with this feature; see Adler 2012, ch. 5). Let  $u^*(\cdot) = au(\cdot)$ , with  $a$  positive. Then note that

$(\prod_{i=1}^N u^*(\mathbf{B}_i^x))^{1/\gamma} = (\prod_{i=1}^N [au(\mathbf{B}_i^x)]^{1/\gamma})^{1/\gamma} = a^{1/\gamma} (\prod_{i=1}^N u(\mathbf{B}_i^x))^{1/\gamma}$ . Thus  $e(x)$  using  $u^*(\cdot)$  as the utility function is a constant factor ( $a^{1-\gamma}$ ) times  $e(x)$  using  $u(\cdot)$  as the utility function, and the ordering of outcomes is therefore the same with either utility function.

How should we arrive at a utility function unique up to a ratio transformation? Let  $\mathbf{B}^{\text{Zero}}$  be a threshold bundle, such that everyone (given their common personal preferences) is indifferent between this bundle and nonexistence. Then let  $u(\cdot)$  be a utility function which (1) is a vNM utility function that represents the common ranking of bundle lotteries, *and* (2) assigns

the number zero to  $\mathbf{B}^{\text{Zero}}$ . Note that if any other  $u^*(.)$  has these properties, there must be a positive  $a$  such that  $u^*(.) = au(.)$ . (Since  $u^*(.)$  is a vNM function, there must be a positive  $a$  and a  $b$  such that  $u^*(.) = au(.) + b$ ; but if  $u^*(\mathbf{B}^{\text{Zero}})$  and  $u(\mathbf{B}^{\text{Zero}})$  are zero it follows that  $b = 0$ ). Now,  $u(.)$  and any  $u^*(.)$  equaling  $au(.)$  can be used as inputs to the isoelastic SWF. On these issues, see generally Adler (2012, ch. 3).

Because utilities must be positive for well-defined and well-behaved distributional weights (see above), all bundles in the outcomes under consideration must be strictly preferred to  $\mathbf{B}^{\text{Zero}}$  and thus assigned positive values by  $u(.)$  and all ratio transformations.

Finally, to see how linear and ratio transformations of utility affect isoelastic distributional weights, recall the formula for the isoelastic distributional weight for individual  $i$  using utility function  $u(.)$ :  $u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$ . Let  $u^*(.) = au(.) + b$ , with  $a$  positive.

Then the formula becomes  $u^*(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) \frac{u^*(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} [au(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) + b] a \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$ .

These weights are a constant multiple of the weights using  $u(.)$ —and thus CBA using them achieves the same ranking of outcomes—iff  $b = 0$ , in which case

$$u^*(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) \frac{u^*(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} [au(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)] a \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c} a^1 u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s) \frac{u(c_i^s, \mathbf{p}^s, \mathbf{a}_i^s)}{c}$$

If  $u(.)$  is scaled up by positive  $a$ , distributional weights are scaled up (without affecting the ranking of outcomes) by  $a^{1-\gamma}$ .

### C. Simplifications of $MU_i$ : Invariance and CRRA

As above in the discussion of utilitarian distributional weights and invariance, we suppress reference to the price vector, which is fixed at the status quo level. Assume that the common preferences over consumption lotteries, holding fixed nonconsumption attributes  $\mathbf{a}$ , are invariant to the level of  $\mathbf{a}$  and captured by a utility function  $h(.)$  of consumption. Then if  $u(.)$  is a vNM function representing the common preferences,  $u(c, \mathbf{a}) = m(\mathbf{a})h(c) + k(\mathbf{a})$ . Let  $\mathbf{B}^{\text{Zero}} = (c^{\text{Zero}}, \mathbf{a}^{\text{Zero}})$  be a threshold bundle. Then  $m(.)$ ,  $k(.)$  and  $h(.)$  must also be such that  $u(c^{\text{Zero}}, \mathbf{a}^{\text{Zero}}) = m(\mathbf{a}^{\text{Zero}})h(c^{\text{Zero}}) + k(\mathbf{a}^{\text{Zero}}) = 0$ . As above,  $h(.)$  can but need not take the CRRA form. In either event, the formula for isoelastic distributional weights,  $u(c_i^s, \mathbf{a}_i^s) \frac{u(c_i^s, \mathbf{a}_i^s)}{c}$ , simplifies to

$$(m(\mathbf{a}_i^s)h(c_i^s) + k(\mathbf{a}_i^s))^{-\gamma} m(\mathbf{a}_i^s) h'(c_i^s), \text{ which with the CRRA form for } h(.) \text{ becomes } (m(\mathbf{a}_i^s)h(c_i^s) + k(\mathbf{a}_i^s))^{-\gamma} m(\mathbf{a}_i^s)(c_i^s)^{-\lambda}.$$

Note that  $h(.)$  itself may well be negatively valued, e.g., with the CRRA form and  $\lambda \geq 1$ . But if all bundles are above the threshold,  $k(.)$  and  $m(.)$  will be such that  $u(.)$  values are positive.

Consider the additively separable form for  $u(\cdot)$ :  $u(c, \mathbf{a}) = h(c) + k(\mathbf{a})$ . In the case of utilitarian distributional weights, as shown above, this form means that distributional weights are just a function of status quo consumption.  $MU_i = \frac{u(c_i^s, \mathbf{a}_i^s)}{c} = h'(c_i^s)$ . By contrast, even with the additively separable form, isoelastic distributional weights are *not* just a function of status quo consumption.  $u(c_i^s, \mathbf{a}_i^s) \frac{u(c_i^s, \mathbf{a}_i^s)}{c}$  simplifies to  $(h(c_i^s) + k(\mathbf{a}_i^s))^{-\gamma} h'(c_i^s)$ .

Finally, consider  $u^*(\cdot)$  which is a positive ratio transformation of  $u(\cdot)$ , i.e.,  $u^*(\cdot) = au(\cdot)$ , with  $a$  positive. Note that  $u^*(c, \mathbf{a}) = am(\mathbf{a})h(c) + ak(\mathbf{a})$ , and that  $u^*(c^{\text{Zero}}, \mathbf{a}^{\text{Zero}}) = 0$ . Thus the simplified forms just described carry over to  $u^*(\cdot)$ , with distributional weights scaled up by the constant factor  $a^{1-\gamma}$ .

### III. Utilitarian and Isoelastic Distributional Weights under Uncertainty

Assume a finite set  $\mathbf{Z}$  of “states of nature,” with each state  $z$  having probability  $\pi(z)$ . As in the canonical set-up deriving from Savage (1972), a given choice (whether the status quo policy  $s$  of inaction, or the choice of some alternative policy  $P$ ) is a mapping from states to outcomes. More specifically, let  $\mathbf{B}_i^{s,z}$  be the bundle of individual  $i$  in state  $z$ , given the status quo choice; and  $\mathbf{B}_i^{P,z}$  be the bundle of individual  $i$  in state  $z$ , given the choice of policy  $P$ .

Under uncertainty, the utilitarian SWF ranks policies using the rule: policy  $P$  is at least as good as policy  $P^*$  iff 
$$\sum_{z \in \mathbf{Z}} \sum_{i=1}^N u(\mathbf{B}_i^{P,z}) \geq \sum_{z \in \mathbf{Z}} \sum_{i=1}^N u(\mathbf{B}_i^{P^*,z})$$
. Note that the formula

$$\sum_{z \in \mathbf{Z}} \sum_{i=1}^N u(\mathbf{B}_i^{P,z})$$
 is an “ex post” formula, taking the expectation of utilitarian social welfare.

However, it is mathematically equivalent to the “ex ante” formula, summing individuals’

expected utilities: 
$$\sum_{i=1}^N \sum_{z \in \mathbf{Z}} \pi(z) u(\mathbf{B}_i^{P,z})$$
.

By contrast, with a nonutilitarian SWF such as the isoelastic SWF, the “ex post” and “ex ante” formulas are not equivalent. The “ex post” isoelastic SWF ranks policies using the rule: policy  $P$  is at least as good as policy  $P^*$  iff

$$\sum_{z \in \mathbf{Z}} \left( \sum_{i=1}^N u(\mathbf{B}_i^{P,z}) \right)^{\frac{1}{\gamma}} \geq \sum_{z \in \mathbf{Z}} \left( \sum_{i=1}^N u(\mathbf{B}_i^{P^*,z}) \right)^{\frac{1}{\gamma}}$$
. The “ex ante” isoelastic SWF

ranks policies using the rule:  $P$  at least as good as  $P^*$  iff

$$\left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} \pi(z) u(\mathbf{B}_i^{P,z}) \right)^{\frac{1}{\gamma}} \geq \left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} \pi(z) u(\mathbf{B}_i^{P^*,z}) \right)^{\frac{1}{\gamma}}$$
. Note that the term

$(z)u(\mathbf{B}_i^{P,z})$  is individual  $i$ 's expected utility given policy  $P$ , and that the “ex ante” isoelastic approach sums these expected utilities transformed by raising them to the power  $(1-\gamma)$  and multiplying by  $(1-\gamma)^{-1}$ .

Under uncertainty, monetary equivalents can be defined as follows: the change in an individual's consumption which counterbalances the difference in his expected utility between two choices, if the change occurs in every state. More precisely, individual  $i$ 's equivalent variation for policy  $P$  is the amount  $\Delta c_i^P$  such that:

$$(z)u(c_i^{s,z}, c_i^P, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) = (z)u(c_i^{P,z}, \mathbf{p}^{P,z}, \mathbf{a}_i^{P,z}).$$

The construction of utilitarian distributional weights proceeds parallel to the certainty case. The utilitarian ranking of policies is equivalent to that achieved by using the following

formula to rank policies:  $\sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{P,z}, \mathbf{p}^{P,z}, \mathbf{a}_i^{P,z}) - \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z})$ , which simply

subtracts a constant (expected utilitarian social welfare in the status quo) from each side. By definition of the equivalent variation, the formula just provided is in turn equivalent to

$$\sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, c_i^P, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) - \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}).$$

And this can now be

approximated, with the total-differential approximation, as

$$\sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z) \frac{u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z})}{c} \Delta c_i^P - \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z) \frac{u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z})}{c}.$$

Note that the term in brackets is the status quo expected marginal utility of consumption for individual  $i$  (“EMU $_i$ ”). This term sums the marginal utility of consumption in each state—given the bundle of attributes in this state that the individual would possess if the status quo policy  $s$  were chosen—discounted by the probability of the state.

In short, the utilitarian ranking of policies can be approximated by the sum of  $\text{EMU}_i \times \Delta c_i^P$ .

“Ex ante” isoelastic distributional weights are constructed as follows. The ex ante isoelastic ranking of policies is equivalent to that achieved by using the following formula:

$$\left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{P,z}, \mathbf{p}^{P,z}, \mathbf{a}_i^{P,z}) \right)^{\frac{1}{1-\gamma}} - \left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) \right)^{\frac{1}{1-\gamma}}.$$

Substituting

in the equivalent variation, this becomes:

$$\left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, c_i^P, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) \right)^{\frac{1}{1-\gamma}} - \left( \sum_{i=1}^N \sum_{z \in \mathbf{Z}} (z)u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) \right)^{\frac{1}{1-\gamma}}.$$

And

using the total-differential approximation (with the chain rule), this is in turn approximately equal to:

$$\sum_{i=1}^N \sum_{z \in Z} (z) u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) \frac{u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z})}{c_i^z} c_i^P \quad \text{Note that individual } i\text{'s equivalent}$$

variation is now multiplied by the  $EMU_i$  term plus an additional term, to the left, which is the status quo “marginal moral value of expected utility” ( $MMVEU_i$ ): status quo expected utility raised to the power  $-\gamma$ .

It turns out to be substantially more complicated to arrive at distributional weights that approximate the “ex post” isoelastic SWF. Why? If there are  $Z$  states in total, let  $\mathbf{C}$  be a grand vector with  $Z \times N$  entries describing each individual’s state-dependent consumption; and  $\mathbf{C}^s$  be the specific vector of  $Z \times N$  state-dependent consumptions given the choice of the status quo. Let  $\mathbf{A}$  be a similar such grand vector describing each individual’s state-dependent non-consumption attributes, with  $\mathbf{A}^s$  the vector of non-consumption attributes given the status quo.  $\mathbf{P}$  and  $\mathbf{P}^s$  provide the same information regarding prices. Let  $F(\cdot)$  be a continuously differentiable function of the form  $F(\mathbf{C}, \mathbf{P}, \mathbf{A})$ . Finally, let  $\Delta\mathbf{C}$  be a vector of  $N$  state-independent consumption changes ( $\Delta c_1, \dots, \Delta c_N$ ), and let  $\mathbf{C} + \Delta\mathbf{C}$  denote vector  $\mathbf{C}$  with each individual  $i$ ’s consumption increased (in all  $Z$  states) by the  $i$ th element of  $\Delta\mathbf{C}$ , that is,  $\Delta c_i$ .

Reviewing the analysis above, it can be seen that the value assigned by both the utilitarian and “ex ante” isoelastic SWFs to a given policy  $P$  can be expressed in the form  $F(\mathbf{C}^s + \Delta\mathbf{C}^P, \mathbf{P}^s, \mathbf{A}^s)$ , with  $\Delta\mathbf{C}^P$  just equal to  $(\Delta c_1^P, \dots, \Delta c_N^P)$ . (The probabilities of the various states, which are fixed, are reflected in  $F$ .) The total-differential approximation is therefore available in the following form: the difference between the  $F$  value of the policy  $P$  and the  $F$  value of the status quo is approximately  $\sum_{i=1}^N \frac{F}{c_i^z}(\mathbf{C}^s, \mathbf{P}^s, \mathbf{A}^s) c_i^P$ . Note that the partial derivative terms in

this summation depend *just* on each individual’s state-dependent *status quo* consumption and other attributes, plus state-dependent *status quo* prices—the information contained in  $(\mathbf{C}^s, \mathbf{P}^s, \mathbf{A}^s)$ —and that this information, together with a *single* equivalent variation assigned to each individual ( $\Delta c_i^P$ ), can be used to approximate the value of any given policy. The formulas above for using distributional weights to approximate the utilitarian and “ex ante” isoelastic SWFs under uncertainty are specific versions of this general formula.

By contrast, the value assigned to a given policy by the “ex post” isoelastic SWF *cannot* be expressed in the form  $F(\mathbf{C}^s + \Delta\mathbf{C}^P, \mathbf{P}^s, \mathbf{A}^s)$ . It is quite possible for two policies,  $P$  and  $P^*$ , to be such that  $\Delta c_i^P = \Delta c_i^{P^*}$  for all individuals, and yet for the ex post isoelastic SWF to assign the two a different value. For a very simple example, use the isoelastic SWF with  $\gamma = 1/2$ . Assume that there are two states of nature,  $z$  and  $z^+$ , each with probability 0.5; that society is choosing between the status quo, policy  $P$  and policy  $P^*$ ; and that individual utility is just a function of

consumption. Every individual *except* individual  $j$  is unaffected by the choice between the status quo,  $P$ , and  $P^*$ , i.e., each such individual's state-specific consumption is the same with all three.

With the status quo, individual  $j$  has consumption 50 in both states; with policy  $P$ , her consumption is 70 in one state and 40 in another; with policy  $P^*$ , her consumption is 100 in one state and 10 in another. Assume that utility (unique up to a ratio transformation) is a constant multiple  $a$  times consumption. Recall also that the isoelastic SWF is invariant to ratio transformations of utility, so we can just set  $a = 1$ . Note now that individual  $j$ 's  $\Delta c_j^P$  value is 5 units of consumption, and that her  $\Delta c_j^{P^*}$  value is also 5. (Given that utility is a constant multiple of consumption, or more generally linear in consumption, she is indifferent between receiving 55 units of consumption for sure, a gamble with equal probabilities of 70 and 40, and a gamble with equal probabilities of 100 and 10.)

However, a quick calculation will show that the ex post isoelastic SWF with  $\gamma = 1/2$  assigns the status quo a value of 14.14 (plus a constant  $C$  to reflect the consumption of individuals other than  $j$ ); it assigns policy  $P$  a value of 14.69 (plus the same  $C$ ); and it assigns policy  $P^*$  a value of 13.16 (plus the same  $C$ ). Thus  $P$  and  $P^*$  are assigned different values even though  $\Delta c_j^P = \Delta c_j^{P^*}$  and  $\Delta c_i^P = \Delta c_i^{P^*}$  for all other individuals  $i$ . Indeed, note that the ex post isoelastic SWF prefers  $P$  to the status quo but prefers the status quo to  $P^*$ ; although the two policies correspond to the very same vectors of equivalent variations, one is seen by the ex post approach as better than the status quo, another worse.<sup>12</sup>

The ranking of policies by the ex post isoelastic SWF *can* be mirrored (precisely, not just approximately), by CBA with distributional weights that are policy-specific, i.e., depend upon state-dependent prices and individuals' state-dependent consumption and other attributes for each policy (not merely the status quo). Let the policy specific weight for individual  $i$ ,  $w_i^P$ , be defined as follows (assuming equivalent variations are nonzero).

$w_i^P = \frac{(z)u(\mathbf{B}_i^{P,z})^1}{(z)u(\mathbf{B}_i^{s,z})^1} / c_i^P$ . Then it follows that the ex post

approach ranks  $P$  at least as good as  $P^*$  iff  $\sum_{i=1}^N w_i^P c_i^P \geq \sum_{i=1}^N w_i^{P^*} c_i^{P^*}$ . However, in this case the

calculation of distributional weights is more laborious than directly applying the ex post formula; it's therefore hard to see why policymakers wouldn't simply do the latter.

The ranking of policies by the ex post isoelastic SWF can also be approximated by assigning each individual a  $Z$ -entry vector of state-specific equivalent variations for each policy  $P$  (rather than a single equivalent variation  $\Delta c_i^P$ ), and combining this with information about status quo prices and individuals' status quo consumption and other attributes. Let  $\Delta c_i^{P,z}$  be

<sup>12</sup> The possibility of such a case, in turn, reflects the fact that the ex post isoelastic SWF violates the ex ante Pareto principle. See generally Adler (2012, ch. 7).

such that  $u(c_i^{s,z} + \Delta c_i^{P,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) = u(c_i^{P,z}, \mathbf{p}^{P,z}, \mathbf{a}_i^{P,z})$ . Then, using the total differential approximation, the ex post isoelastic ranking of policies *can* be approximated by assigning each

$P$  the following value: 
$$u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z}) - \frac{u(c_i^{s,z}, \mathbf{p}^{s,z}, \mathbf{a}_i^{s,z})}{c_i^{P,z}} \Delta c_i^{P,z}.$$
 But determining an

individual array of state-specific equivalent variations could be quite cumbersome, and it is hard to see why this approximation offers much economy in decisional effort over direct application of the ex post isoelastic formula.

#### IV. The Value of Statistical Life

VSL is the marginal rate of substitution between survival probability and consumption/income/wealth. The simple, standard, one-period model of VSL, relied upon in the text, works as follows. Each individual in the status quo has a consumption amount  $c_i$  and a probability  $p_i$  of surviving the period.<sup>13</sup> A policy changes these consumption amounts and/or probabilities. The utility of consumption is structured as follows. Let  $h_{survive}(c) = u(c, survive)$  be the utility of consumption  $c$  if the individual survives the period, and  $h_{die}(c) = u(c, die)$  be the utility of consumption  $c$  if the individual does not survive. The standard model assumes that, for all  $c$ :  $h_{survive}(c) > h_{die}(c)$ ;  $h'_{survive}(c) > h'_{die}(c) \geq 0$ ;  $h''_{survive}(c) \leq 0$ ,  $h''_{die}(c) \leq 0$ .

For convenience, I omit the superscript “ $s$ ”; unless otherwise noted, consumption amounts and probabilities are status quo values.

Let  $U_i$  be individual  $i$ 's status quo expected utility, i.e.,  $p_i h_{survive}(c_i) + (1-p_i) h_{die}(c_i)$ . VSL <sub>$i$</sub>  is  $\frac{U_i}{p_i} \bigg/ \frac{U_i}{c_i}$ , which equals  $\frac{h_{survive}(c_i) - h_{die}(c_i)}{p_i h_{survive}(c_i) - (1-p_i) h_{die}(c_i)}$ .

Consider a policy  $P$  which changes individual  $i$ 's survival probability by  $\Delta p_i^P$  and his consumption by  $\Delta y_i^P$ . Let  $U(c, p)$  denote expected utility as function of consumption and survival probability.<sup>14</sup> Individual  $i$ 's equivalent variation for this policy (an equivalent variation under uncertainty, in the sense discussed in the previous part of this Appendix), is the amount  $\Delta c_i^P$  such that  $U(c_i + \Delta c_i^P, p_i) = U(c_i + \Delta y_i^P, p_i + \Delta p_i^P)$ . Applying the total-differential approximation to both sides of this equation, we have that

<sup>13</sup> Consumption in this simple model is state-independent. In the status quo, an individual has the same consumption in all states of nature; a policy may change his consumption to a different, single, level in all states.

<sup>14</sup> The functions  $\frac{U_i}{c_i}$  and  $\frac{U_i}{p_i}$  are, of course, shorthands for  $\frac{U}{c}(c, p)$  and  $\frac{U}{p}(c, p)$  with the  $c$  input understood to be some level of  $c_i$ , and the  $p$  input some level of  $p_i$ .

$$\frac{U_i}{c_i} c_i^P - \frac{U_i}{c_i} y_i^P - \frac{U_i}{p_i} p_i^P, \text{ i.e., } \Delta c_i^P \approx \Delta y_i^P + \text{VSL}_i \Delta p_i^P. \text{ The equivalent variation is}$$

approximately equal to the consumption change plus VSL times the survival probability change; in other words, VSL is the rate at which small probability changes are converted into equivalent consumption changes.

My analysis in the text adds two simplifying assumptions:  $h_{\text{survive}}(\cdot)$  is CRRA, and  $h_{\text{die}}(\cdot)$  equals zero. (This latter assumption, of course, means, that  $h_{\text{die}}(c)$  is the same value for all consumption amounts.) I also assume that there exists a subsistence level  $c^*$  of consumption such that  $h_{\text{survive}}(c^*) = h_{\text{die}}(c)$  for any  $c$ .

$u(\cdot)$  and thus  $h_{\text{survive}}(\cdot)$  and  $h_{\text{die}}(\cdot)$  are unique up to a linear transformation, which suffices for purposes of VSL (or utilitarian distributional weights), but not isoelastic weights. However, as discussed earlier, the threshold bundle used to construct a utility function unique up to a ratio transformation is, plausibly, the bundle  $\mathbf{B}^{\text{zero}}$  such that everyone (given common preferences) is indifferent between  $\mathbf{B}^{\text{zero}}$  and nonexistence. And the model of VSL now at hand already defines that bundle, namely,  $(c^*, \text{survive})$  such that  $u(c^*, \text{survive}) = u(c, \text{die})$  for any  $c$ .

We can therefore rescale  $u(\cdot)$  as follows. Let  $v(\cdot) = u(\cdot) - u(c^*, \text{survive})$ . Thus  $v(c^*, \text{survive}) = 0 = v(c, \text{die})$  for any  $c$ . Adding in that  $h_{\text{survive}}(\cdot) = u(\cdot, \text{survive})$  is CRRA, we have that  $u(c, \text{survive}) = (1 - \lambda)^{-1} (c)^{1-\lambda}$ , and thus  $v(c, \text{survive})$  equals:  $(1 - \lambda)^{-1} (c)^{1-\lambda} - (1 - \lambda)^{-1} (c^*)^{1-\lambda}$ .

The formula for  $\text{VSL}_i$  simplifies to:  $\frac{(1 - \lambda)^{-1} (c_i^{1-\lambda} - (c^*)^{1-\lambda})}{p_i c_i}$ . It is this formula which is used to

calculate the VSL values in Table 6.

Let  $V(c_i, p_i)$  be the expected  $v(\cdot)$  value as a function of consumption and survival probability. Recall that  $\Delta p_i^P$  and  $\Delta y_i^P$  are the changes in individual  $i$ 's survival probability and consumption, respectively, produced by policy  $P$ . The utilitarian SWF ranks policies with the formula  $\sum_{i=1}^N V(c_i, y_i^P, p_i, p_i^P)$ , which is equivalent to ranking them using the formula

$$\sum_{i=1}^N V(c_i, y_i^P, p_i, p_i^P) - \sum_{i=1}^N V(c_i, p_i) . \text{ By the total differential approximation, this is}$$

approximately

Recall now the general formula for utilitarian distributional weights under uncertainty: the sum of equivalent variations multiplied by  $EMU_i$ . In the context at hand,  $EMU_i$  is precisely  $\frac{V_i}{c_i}$  which equals  $p_i c_i^{-\lambda}$  given the CRRA assumption.

In short, with the CRRA assumption, the utilitarian SWF is approximated by  $\prod_{i=1}^N p_i c_i VSL_i p_i^P y_i^P$ . This can be seen as an instance of the general formula for utilitarian weights under uncertainty, with the equivalent variation in turn approximated as  $\Delta y_i^P + VSL_i \Delta p_i^P$ .

Similarly, the ex ante isoelastic ranking of policies is given by

$$(1) \prod_{i=1}^N V(c_i, y_i^P, p_i, p_i^P)^1, \text{ or equivalently by}$$

$$(1) \prod_{i=1}^N V(c_i, y_i^P, p_i, p_i^P)^1 = (1) \prod_{i=1}^N V(c_i, p_i)^1. \text{ This is approximately}$$

$$\prod_{i=1}^N V(c_i, p_i) \left( \frac{V_i}{c_i} y_i^P - \frac{V_i}{p_i} p_i^P \right), \text{ which becomes } \prod_{i=1}^N V(c_i, p_i) p_i c_i (VSL_i p_i^P y_i^P), \text{ with}$$

$V(c_i, p_i) = p_i ((1-\lambda)^{-1} (c_i)^{1-\lambda} - (1-\lambda)^{-1} (c^*)^{1-\lambda})$ . This is an instance of the general formula earlier for the ex ante isoelastic SWF under uncertainty, with  $MMVEU_i$  here  $V(c_i, p_i)^{-\gamma}$ .

## V. Taxation

Here I provide a brief argument for why a legislature that engages in taxation to maximize a utilitarian or isoelastic SWF will redistribute consumption so that distributional weights in light of that SWF are identical for all individuals—under the first-best conditions of a lump-sum tax that can be administered without administrative costs, without changing the total stock of each marketed good or individuals' nonconsumption attributes, and with individuals' attributes observable to the taxing authority. As throughout this Appendix, individuals are assumed to have common personal preferences representable by a common utility function.

Assume first, for simplicity, that there is a single consumption good, and that  $c_i$  denotes individual  $i$ 's consumption of that good. Utility takes the form  $u(c_i, \mathbf{a}_i)$ , without prices as an argument, since the individual directly consumes  $c_i$  rather than expending it among an array of goods. (As throughout,  $\mathbf{a}_i$  denotes individual  $i$ 's nonconsumption attributes, including labor supply, health, etc., here assumed to be exogenous to the tax scheme.) Let  $C$  denote the total stock of the good (the total available for distribution to the population after government expenditure on public goods or other programs). The utilitarian legislature maximizes

$\sum_{i=1}^N u(c_i, \mathbf{a}_i)$  subject to the constraint that  $\sum_{i=1}^N c_i \leq C$ , while the isoelastic legislature maximizes  $(\sum_{i=1}^N u(c_i, \mathbf{a}_i))^{\lambda}$ , subject to the same constraint. Straightforward constrained optimization shows that, at the utilitarian optimum, for any pair of individuals  $i$  and  $j$ ,  $\frac{u}{c}(c_i, \mathbf{a}_i) = \frac{u}{c}(c_j, \mathbf{a}_j)$ . Similarly, at the isoelastic optimum,  $u(c_i, \mathbf{a}_i) = u(c_j, \mathbf{a}_j) = \frac{u}{c}(c_j, \mathbf{a}_j)$ .

Consider now the case in which there are  $M$  types of marketed goods,  $m=1$  to  $M$ . Let  $g_i^m$  denote individual  $i$ 's holding of good  $m$ , and  $\mathbf{g}_i$  her vector of all  $M$  marketed goods. The direct utility function,  $v(\cdot)$ , takes the form  $v(\mathbf{g}_i, \mathbf{a}_i)$ . The total stock of each marketed good  $m$  is fixed at  $S_m$ . The legislature maximizes a utilitarian SWF or isoelastic SWF —  $\sum_{i=1}^N v(\mathbf{g}_i, \mathbf{a}_i)$  or

$(\sum_{i=1}^N v(\mathbf{g}_i, \mathbf{a}_i))^{\lambda}$ , respectively — subject to  $M$  constraints, each of the form  $\sum_{i=1}^N g_i^m = S_m$ . In

the utilitarian case, the Lagrangian for the optimal allocation of the  $M$  goods across the  $N$  individuals takes the form:  $\sum_{i=1}^N v(\mathbf{g}_i, \mathbf{a}_i) - \lambda_1 (\sum_{i=1}^N g_i^1 - S_1) - \dots - \lambda_M (\sum_{i=1}^N g_i^M - S_M)$ . It follows that the utilitarian optimal allocation is such that, for each good  $m$ , and any two individuals  $i$  and  $j$ ,

$\frac{v}{g^m}(\mathbf{g}_i, \mathbf{a}_i) = \frac{v}{g^m}(\mathbf{g}_j, \mathbf{a}_j)$ . Similarly, in the isoelastic case, the optimal allocation is such that

$v(\mathbf{g}_i, \mathbf{a}_i) = \frac{v}{g^m}(\mathbf{g}_i, \mathbf{a}_i) = v(\mathbf{g}_j, \mathbf{a}_j) = \frac{v}{g^m}(\mathbf{g}_j, \mathbf{a}_j)$  for any two individuals  $i$  and  $j$ .

This analysis shows that the marginal utilitarian or isoelastic social value of *each good* is equalized across individuals at the optimum. In turn, the second welfare theorem shows that the utilitarian optimal allocation is the market equilibrium of some set of prices  $\mathbf{p}$  and lump sum consumption (expenditure) amount  $c_i$  for each individual,<sup>15</sup> with  $\sum_{m=1}^M p^m g_i^m = c_i$ ; and the same is true for the isoelastic optimum. At each optimum (whether utilitarian or isoelastic), each

<sup>15</sup> On the second welfare theorem, see Mas-Collel, Whinston and Green (1995, pp. 551-58).

For the second welfare theorem to obtain, we must assume that preferences have an appropriate structure. Also, the second theorem is normally stated for the case in which individual utility is merely a function of marketed goods. However, individuals *are* allowed to have heterogeneous preferences over goods.  $v(\cdot)$  corresponds to a conditional utility function over good bundles for each possible individual array of nonconsumption attributes. In effect, individuals with attributes  $\mathbf{a}$  have certain preferences over goods, individuals with attributes  $\mathbf{a}^*$  possibly different preferences, etc. The second welfare theorem then says that *any* desired allocation of the goods among these individuals (if Pareto optimal) can be achieved as a market equilibrium.

individual  $i$  maximizes her utility given that her total expenditure is constrained to be  $c_i$ . With  $u(\cdot)$  the indirect utility function,  $u(c_i, \mathbf{p}, \mathbf{a}_i) = v(\mathbf{g}_i, \mathbf{a}_i)$ , and also  $\frac{u}{c}(c_i, \mathbf{p}, \mathbf{a}_i) = (1/p^m) \frac{v}{g^m}(\mathbf{g}_i, \mathbf{a}_i)$  for every one of the  $M$  goods.

But  $\frac{v}{g^m}(\mathbf{g}_i, \mathbf{a}_i) = \frac{v}{g^m}(\mathbf{g}_j, \mathbf{a}_j)$  for any two individuals  $i$  and  $j$  at the utilitarian optimum.

It therefore follows that  $\frac{u}{c}(c_i, \mathbf{p}, \mathbf{a}_i) = \frac{u}{c}(c_j, \mathbf{p}, \mathbf{a}_j)$  for any two individuals  $i$  and  $j$  at the

utilitarian optimum. Similarly,  $v(\mathbf{g}_i, \mathbf{a}_i) = \frac{v}{g^m}(\mathbf{g}_i, \mathbf{a}_i) \cdot g^m(\mathbf{g}_i, \mathbf{a}_i) = v(\mathbf{g}_j, \mathbf{a}_j) = \frac{v}{g^m}(\mathbf{g}_j, \mathbf{a}_j) \cdot g^m(\mathbf{g}_j, \mathbf{a}_j)$  for any two individuals at the isoelastic optimum. Thus

$u(c_i, \mathbf{p}, \mathbf{a}_i) = \frac{u}{c}(c_i, \mathbf{p}, \mathbf{a}_i) \cdot c_i = \frac{u}{c}(c_j, \mathbf{p}, \mathbf{a}_j) \cdot c_j = u(c_j, \mathbf{p}, \mathbf{a}_j)$  at the isoelastic optimum.