EDITED BY

MICHAEL
MORAN

MARTIN
REIN

ROBERT E.
GOODIN

The Oxford Handbook *of*
PUBLIC POLICY

Shortcomings in economic assumptions notwithstanding, economic evaluations in health care provision are more in demand than ever before, greatly spurred by the ever-growing share of GDP that is absorbed by the treatment of nations' aging populations. Carrying out CBAs in such policy contexts promises guidance for decision makers as to the optimal distribution of medical manpower, R&D funding, reimbursement practices, capital controls, and safety regulations. Costs and benefits accrue at three different points, or channels where health care is provided: cure (to improve health), care (to retain dignity for those who are sick), and prevention (to reduce the probability of illness or premature death). The benefits in these channels are established by valuing the respective effects a policy has on the state of health of the individual(s) in question. The methods used to conduct this activity have attracted their own set of criticisms. They are similar to the charges elucidated in Section 3 above and will therefore not be rehearsed here.

Rather, we direct our attention to a related issue, the aggregation of attributes of well-being, which represents itself as soon as health improvements *have been* valued. Aggregation is a task not confined to health care but is pursued in all policy domains and for all goods and services that governments provide. Aggregation needs to be done over different outcomes of varied interventions undertaken on different problems. Staying with health care as a policy domain, for life-threatening diseases such as coronary bypass surgery or tetanus the primary outcome will obviously be defined as death or survival. Case fatality rate and survival rate may in such cases be good indicators of the achievements of heath care reached. Each survival can then be indexed with the value 1 and each fatality with 0. Treatment of most other illnesses—or for that matter, effects of other policy decisions on well-being—does not result in such binary outcomes, however, and measuring them in such a way means that everyone who survives a medical intervention is given the same value, no matter whether the person is confined to bed or is actively able to play sports. A more accurate measure would be required for these cases, one that is able to capture benefits in the form of subsequent *grades* of well-being between the two end points of the spectrum.

In a move to derive a methodology suitable to develop such an index, scholars began from the 1970s onwards, to define health in terms of "utility of life" (Torrance, Thomas, and Sackett 1972; Zeckhauser and Shephard 1976). Three decades of research and numerous refinements later, utility of life has come to be calculated along two dimensions: (*a*) the duration of life as measured in life years and (*b*) the quality of life as experienced by the individual's physical, social, and emotional functioning. The latter is elicited via patient questionnaires and interviews, where rating scale, time trade-off, or standard gambling techniques (of which more will be heard in a moment) are applied across a multitude of domains—including mobility, emotion, cognition, and pain—so as to arrive at the weighted preference that each domain commands (Drummond et al. 1997, 150–83). The greater the preference for a particular health state, the greater the "utility" associated with it. Utilities of health states are generally expressed on a numerical scale ranging from 0 to 1, in which 0 represents the utility of the state "dead" and 1 the utility of a state lived in "perfect health." Finally, utilities are multiplied by the remainder of an individual's lifetime

for each outcome to calculate so-called "quality-adjusted life years" (QALYs). The QALY benefit associated with any given intervention is calculated as the difference between the QALYs available with that intervention and the QALYs available without that intervention. The results can then be used to create "cost-per-QALY" rankings for different interventions which aids in deciding on "best-buy" strategies, and to develop statistics on "disability-adjusted life-year expectancies" (DALYs) across countries (WHO 2000, 176–83; Murray 1996).

The QALYs approach is an exercise in what is commonly called "multi criteria mapping" and thus akin to methods developed to address aggregation issues in other policy domains. It soon established itself as the most sophisticated and therefore default methodology for measuring and aggregating individual levels of human well-being in general and quality of life in health care in particular. In no other policy sector has there been developed a similarly refined approach. And as a non-monetary standard it has the added benefit of bypassing the criticisms about monetary valuation that we elaborated upon in the previous section.

Despite the advantages of using a single indicator to measure the effectiveness of health care interventions, QALYs have been widely criticized on ethical, conceptual, and operational grounds, casting doubts on whether the underlying methodology actually solves the problem of incommensurability. The possibility of combining quantity and quality of life in a single index is rooted in the school of political philosophy known as utilitarianism. It is the foundation for the economic analysis of individual behaviour and emerged in the works of Jeremy Bentham and John Stuart Mill in the eighteenth and nineteenth centuries respectively. Now known as the "interpersonal comparison of well-being" problem, it has kept philosophers on their toes ever since (Elster and Roemer 1991).[3] Bentham's intention was to provide the British Parliament with a political theory that could be used to construct sound and rational policies rather than letting them rely on vague and biased intuitions. The theory's main prescription was to enact laws that are dictated by the principle of utility, when in like manner the tendency which it has to augment the utility (or "happiness" as Bentham called it) of the community is greater than any which it has to diminish (Bentham 1970). In what became later known as classical utilitarianism, this principle directs the policy maker to maximize the utility of the members of a society.

Utilitarian theory has been persistently attractive to generations of policy makers and political theorists because of its simplicity; its scientific allure as a theory that can be written down as a mathematical formula; and its concern for human welfare as the core of moral philosophy. Yet it has also attracted its fair share of criticism, resulting in many authors proposing modifications and redefinitions to make the theory more palatable. This is certainly not the place to rehearse this debate. The reader may refer to the extensive research produced on the topic, with the collection edited, for example, by Glover (1990) providing a good starting point. Sen (1987, 39) is more useful for us in that he has drawn out the elementary requirements of any utilitarian

---

[3] We use "utility" here interchangeably with the terms "welfare" and "well being" as the satisfaction accruing to an individual from the consumption of a good or service.

moral principle. These are (1) *welfarism*, requiring that the goodness of a state of affairs be a function only of the utility information regarding that state; (2) *sum ranking*, requiring that utility information regarding any state be assessed by looking only at the sum total of all the utilities in that state; and (3) *consequentialism*, requiring that every choice, whether of actions, institutions, motivations, rules, etc., be ultimately determined by the goodness of the consequent state of affairs.

Note that the first requirement about welfarism can only be made to work if individuals are assumed to be able to evaluate their utility; if that utility can be made known to interested third parties, such as policy makers through some sort of valuation; and if that valuation can be measured in quantitative terms. These assumptions have already been questioned in Section 3 when we discussed the case of environmental goods. It is the second requirement on sum ranking which we are concerned with in the current context of aggregation of utilities and QALYs. Bentham insisted that sum ranking is possible because, to him, the item to be aggregated (happiness) denoted only one type of experience (the feeling of pleasure). Hence, utility was in his view easily aggregated across lives, for it was only one, not multiple experiences that people would encounter. It didn't take long before philosophers objected that some pleasures differ in kind according to the value individuals attach to them. And these are not the same across lives.[4]

Given the multiplicity of states of health that individuals might experience, the question then remains whether it is possible to know how much healthier some are compared to others. We are certainly able to make such a comparison in an ordinal sense, e.g. I can stipulate that I feel better than someone who is in great physical pain. However, to compare utilities across lives, I need to be able to make the comparison in a cardinal sense, i.e. I need to know exactly *how much better* I am. Cardinality, in turn, implies two requirements that need to be satisfied (Bossert 1991): (1) a number must be attached to the outcome that represents the strength of the preference relative to others, so that a health state of, say, 0.6 is three times better than one of 0.2; and (2) the scale must have an equal interval property where equal differences at different points along the response scale are equally meaningful, so that boosting a patient from, say, 0.1 to 0.2 on that scale is of equal benefit to raising someone from 0.8 to 0.9.

Health scientists and policy makers have recently started to develop various preference elicitation techniques in an effort to calculate the required QALY weightings. Various psychological studies suggest that because of cognitive limitations in humans, the techniques do not always elicit responses that satisfy the two requirements. With the *rating scale* approach, for example, individuals are asked to rank health outcomes from

[4] The utility concept as used by most economists and philosophers in the nineteenth and twentieth centuries is theoretically distinct from the utility used in the QALY methodology. The former describes decisions where goods are received with certainty, whereas the latter does so for probabilistic outcomes under uncertainty. Decision theory under uncertainty aspires to the more rigid requirements as stipulated by the so called von Neumann Morgenstern utility theory (von Neumann and Morgenstern 1947), whereas the conventional philosophical/economic understanding sees a utility merely as the satisfaction of preferences. For our discussion this is no relevant distinction, however: NM utilities cover decision making theory at the individual level only and cannot be used to compare welfare between individuals (Zeckhauser and Schaefer 1975, 41; Drummond et al. 1997, 150).

most preferred to least preferred and to place them on a scale such that the intervals between placements correspond to the differences in preference as perceived by the individual. However, psychologists have challenged the meaningfulness of the cardinal statements thus produced by respondents. As Bleichrodt and Johannesson (1997) argue, subjective impressions cannot be discriminated equally at each level of a scale. Individuals will attempt to use categories equally often and spread their responses when cases are actually close together (the "spacing out" bias), or they compress them when the underlying attributes are actually far apart (the "end-of-scale" bias).

The standard gamble, as a second method, induces the individual to choose between two alternatives: (*a*) no treatment at all which will result in a specified state of ill health, or (*b*) treatment that could result in either death or illness-free health, each with a probability of p and 1−p respectively. The probability is then varied until the respondent is indifferent between the two alternatives, thus producing the preference score sought after. Tversky, Slovic, and Kahneman (1990), however, have shown through various laboratory experiments that individuals have the tendency to reverse previously revealed preferences. They might use inappropriate psychological representations and simplifying heuristics that misdirect their decisions. Psychologists have attributed this phenomenon to the serial way by which individuals process information: they use an anchoring technique for the first piece of information and then gradually adjust their decision making with each additional piece of information they obtain.

Finally, the *time trade-off* presents individuals with a choice of living for a defined amount of time in perfect health or a variable amount of time in an alternative state that is less desirable. The time is varied until the respondent is indifferent between the two alternatives. The method's application, however, has found patients to prefer, for example, immediate death to being in a state of mild dysfunction for three months. This suggests that individuals misunderstand the nature of the trade-off, reducing the meaningfulness of the results on a utility scale that ranges between 0 and 1.

Patients' responses as well as the metric underlying their measurement cannot, then, be standardized across individuals. Epistemological difficulties remain when adding up or comparing subjective levels of satisfaction that the consumption of goods gives to individuals (Nord 1999). The preference elicitation techniques used with the QALY approach encounter too many teething problems that prevent policy makers from uncovering stable and consistent preferences revealing true commensurate valuations. Notably, the failure to make attributes of well-being commensurate does not mean that comparisons are futile exercises. Incommensurability does not deny the possibility of comparisons of course. Neither does it need to be inconsistent with fundamental assumptions in decision theory: reason-guided choice is still possible even without commensurability, as the data underlying QALYs are still useful to make more simple comparisons through ordinal rankings (Sunstein 1997, 39). Yet, they lack the precision that is required to impute them into economic methodologies such as CBA.

More exchange between psychologists, economists, and philosophers seems necessary. For the case of health care in the UK, for example, the National Center for Research Methodology (NCRM) and the National Institute for Clinical Excellence (NICE) have

recently commissioned joint research projects with the aim of determining the societal value of a QALY.[5] This project addresses, among other issues, the conceptual link between a QALY and an individual's WTP as well as the relative value of health gains to different beneficiaries, according to personal attributes such as age, education, and geography. These initiatives could shed more light on the problem at hand. Until solutions are developed from those (and other) findings, however, the second requirement on sum ranking that Sen specified for utilitarian theory remains unsatisfied.

To be sure, as Kymlicka (2001, 18) rightly reminds us, in daily life practical reasoning constantly requires us to make decisions about how to balance different kinds of goods that are incommensurable, by simply judging what is better or worse overall. While we might go along with his assessment for the individual decisions we make in our personal lives, we believe it is an ill-advised position to take for the analysis of public policy. The economic evaluation techniques used to arrive at policy decisions differ in their level of complexity from the balancing acts between the comparatively few personal values that inform our individual choices. We can revisit and reassess the ordinal rankings we have made in a personal choice situation at any given time. Economic evaluation techniques, by contrast, balance many more preferences and values that are held by markedly more individuals and eventually produce only one (usually quantitative) recommendation. From that moment on, they conceal the complex weighing process between the different cardinal attributes that had been imputed beforehand.

Admittedly, for evaluation techniques to work the imputed preferences and values need to be made explicit in the first place, which is an approach preferable to making policy choices on the basis of decision makers' implicit (and therefore concealed) assumptions and preferences. Yet, once all of the relevant goods are aligned along a single metric, they are no longer visible, or perhaps become invisible (Sunstein 1997, 50). People can no longer make judgements based on qualitative differences. Hence, if we want the policy recommendation to be meaningful and accurate we need to ensure that the numerical values imputed into the analysis at the outset have been compared and aggregated accurately. This demonstrably does not always hold true, in which case the policy choice needs to be made through alternative measures. Some of these we will present in Section 6 below.

5. The Intrinsic Value Problem

At the end of Section 3 we introduced the concepts of "existence value," "exchange value," and "use value" to our discussion. We defined existence value as a value that a good can have independent of the effects it produces for human well-being, such as the survival of species. We also contended that exchange value, as the metric that is

imputed into economic evaluations, bears no necessary connection to the value in use that produces the benefit to individuals and thus augments human well-being.

There is a crucial link between these three concepts that merits further exploration: economic evaluations impose a unitary standard (usually money) on the valuation and comparison of goods and thus subordinate both existence value and use value to the new standard of exchange value. While we have already drawn out some aspects of this relationship for *objects* (i.e. environmental goods in Section 3 and health care services in Section 4), we will in this section, develop that point in more detail for *subjects*. We will argue that the intrinsic value of human beings (as the equivalent to the existence value of objects) is crowded out by economic evaluations.

To understand why, let us assume that in some distant future, the problem of valuation and aggregation expounded earlier will have been solved and that it is therefore possible to evaluate policy programs according to the extent to which they maximize benefits to society. Now consider the following simplified case borrowed from Harris (1975): a hospital has admitted four patients who are all bound to die if no suitable organ donor is soon to be found. The next morning, the postman enters the building to deliver his daily load of letters and parcels. From previous conversations the nurse recalls that he would be a suitable donor for all four patients. As a possible route of action she could now kill him, harvest his organs, and thus enable the four patients to survive. If numbers count and we conduct a simple CBA we would have to conclude that sacrificing the postman is the superior alternative: four lives are more valuable than one and the highest aggregate level of welfare is achieved if the postman dies and the four patients live.

Most of us would consider this option as objectionable of course. In most contexts it strikes us intuitively as unfair if a few may be sacrificed for the benefit of the greater good of the many. Yet, given the economic rationale of benefit maximization, it is justifiable, if not mandatory, to proceed that way. The problem we encounter here is caused by the formally equal way by which these evaluations treat human beings: every individual counts as one and can thus be added up to, or traded against somebody else. This observation is akin to the phenomenon of "commodification" originally developed by Karl Marx (1964, 96–105). In capitalist societies, so Marx argued, the mode of production comes under private ownership, commodity production proliferates, and labor division becomes increasingly fragmented. Forced to sell their labor power to survive, workers themselves become akin to a commodity and are reduced from the status of a qualitative individual to mere exchange value in the form of labor. Where once the goal of production was the simple satisfaction of needs, and exchange was driven through the need for the other's use value, capitalism eliminates individual exchange. It subordinates use value to exchange value and establishes exchange value as an independent logic. In the extreme but quite common form of trading stocks, for example, there is no longer a physical referent at all: money is made out of money with no apparent connection to the world of real commodities.

The reduction of human beings to a number—either expressed as a simple unit as in the organ donor case or as a monetary WTP value attached to their preferences—

assumes equivalence between attributes of persons and thus dissolves their qualitative differences into the identity of a single quantitative metric. Such a metric might in general solve the problem of aggregation (how to compare levels of well-being), and the monetary metric as exchange value in particular might solve the problem of exchange (how to trade qualitatively unique goods in equal quantitative ratios), but it transforms subjects into abstract entities that are deprived of their unique characteristics.

One such characteristic is that each individual has intrinsic value: we have an interest in our own continued existence and cannot be used solely as a means for assisting other individuals as ends in themselves. Intrinsic values are non-relational: they are not defined relative to some other human being, species, or object, nor to the benefit it might provide to them. My intrinsic value is the value I have in and of myself, beyond any value I might have as a means to further ends. I am therefore to be respected as a rights bearer proper, as an end in myself. Rights are principles that assign claims or entitlements to someone against someone, and are usually interpreted as "trumping" consequential claims made in the name of welfare (Dworkin 1977). That means that I should never be treated in certain ways, even if the calculation of aggregated individual well-being shows that the action which has these effects would be the most beneficial one overall.

Reducing individuals to a monetary metric might change the way we perceive their value to us. Margaret Jane Radin (1996) illustrates the implications for the trade in "commodities" such as sex, children, and body parts and observes that there are not only willing buyers for such commodities but some desperately poor people are willing sellers, too. To her, this reflects a persistent dilemma in liberal societies: freedom of choice is valued but at the same time, choices ought to be restricted to protect the integrity of what it means to be a person. She views this tension as primarily the result of underlying social and economic inequalities, which need not reflect an irreconcilable conflict in the premisses of liberal democracy but a mere setting of the right priorities in distributive policy choices.

Political philosophy has therefore sought to embed intrinsic value and individual rights into some concept of justice, such as a (neo-)Kantian imperative to treat others fairly or Locke's view that people have the right to be protected against the breaches of their rights by the actions of others. Even utilitarians like Mill have endorsed rights and intrinsic values as a possible strategy to maximize utility. Such a position is known as rule utilitarianism, in contrast to act utilitarianism which is the view Bentham originally suggested. It postulates that the principle of utility can yield a notion of "rights" if we appreciate the way a person's rights are defined by rules regarding the treatment of human beings that are by and large utility maximizing.

This is no place to develop the pros and cons of any of these concepts. It is important to note, however, that while constraining economic evaluations through intrinsic values and individual rights can be attractive to a great variety of traditionally juxtaposed theories of morality, the resulting consensus in political philosophy cannot be transferred easily to public policy formation or economic evaluation techniques. This follows because, to follow Ruth Chang's (1997, 5–23) helpful distinction, intrinsic values give rise to the problem of ordinal *incomparability*.

The reader might recall from Section 4 that we concluded that attributes of well-being are incommensurable across lives, i.e. that they cannot be compared cardinally for the purpose of aggregation, but that at least ordinal comparisons are available as a basis for rational choice. We now encounter the more severe case where the relevant imputations for the analysis are not even comparable in that latter sense.

This follows because the practical role of intrinsic values is neither to prescribe an end to be maximized nor to prescribe an attitude toward an aggregate. As such there are multiple ways in which we can sharpen our understanding of a person's intrinsic value, such as by love, respect, honor, or admiration. In some cases one understanding might be privileged while in another it isn't. This vagueness disallows for any strand of the usual trichotomy of comparison ("better than," "worse than," "equally good as") to hold, which applies to comparisons between intrinsic values themselves as much as between them and other quantifiable values.

While incomparability might be less of a problem for clear-cut cases such as the life-or-death choices to be made in the organ transplant scenario mentioned earlier, other policy decisions are more clearly subject to this limitation. Health care, to stay in the same policy domain, does not only suffer from a lack of organs, for example. Hospital beds, technical equipment, and medical personnel, too, are scarce resources that can be distributed among patients in different ways. Economic evaluations would recommend that these should be used less intensively for the care of acute or incurable patients as they require far more of them than does the care of convalescing patients. Similarly, applying the QALY approach explicated in Section 4 to the optional treatment of either an elderly person or a young child would result in the preference to be given to the latter, because QALY scores are particularly high for those who still have many years to live and therefore have a greater "capacity to benefit." Economic evaluations applied in an unconstrained way would therefore lead to the marginalization of the incurable, chronically ill, or elderly. They would override individuals' intrinsic value in terms of their dignity and possibly, their right to live.

To be sure, in some contexts an intelligible response that bypasses the intrinsic value problem is possible. The application of distributional weights, for example, can go a long way to ensure an equitable distribution of scarce resources that does not neglect groups who are in need (Layard and Walters 2001). However, while the existence of a tangible criterion to define disadvantage allows us to identify some such groups—e.g. income levels as an indicator that demarcates the needy poor from the non-needy rich—other groups which we deem worthy of special consideration, and would ideally want to apply appropriate distributional weights to, are less lucidly identified. How, for example, should we weigh the feelings of love, respect, honor, or admiration by which we grant a person her intrinsic value? How do we gauge the underlying psychological processes? Our choice between these feelings does not proceed on some measurable comparison but on the more intangible principle of obligation.

Intrinsic values cannot be ranked ordinally in a meaningful way then. There is no way to incorporate them into any type of evaluation. The policy maker is thus faced with a situation in which he can choose to either (1) ignore the intrinsic value, or (2) admit it as a constraint and reject the policy recommendation under review.

The former will then judge the recommendation to be permissible whereas on the latter it is impermissible. Judging the policy as impermissible, in turn, implies that any benefits which would result from rights-incompatible actions must be excluded from the action decision altogether. It places limits on what would otherwise be the implication of aggregative economic evaluations and restricts governmental action.

This is, of course, not a satisfying conclusion to arrive at because our following option 2 puts the whole exercise of economic evaluation into question in the first place while under option 1 intrinsic values are crowded out and "forgotten" by the imperative of identifying, collecting, measuring, and aggregating other values that *are* comparable.

Two alternative and somewhat juxtaposed approaches to the dilemma seem to be on offer both of which, however, require further refinement and specification if they are to provide meaningful solutions. There is, first, the suggestion made by Shrader-Frechette (1991, ch. 11) that each group affected by a proposed policy program should conduct their own economic evaluation as an intermediate stage of a more extensive process of participative justice. This approach would not only allow for a separate assessment of intrinsic values and a weighing of their merits. It would also reflect different methodological, ethical, and social assumptions and thus portray all sides of a given story. The end result would then be likely to be an evaluation with a multidimensional array of benefits and costs. Alternatively, we might want to embrace the work begun by Scanlon (1991) on the compatibility of the ethical and economic conception of value that individuals attach to human well-being. Instead of requiring various stakeholder groups to carry out multiple evaluations that are later democratically deliberated upon, Scanlon suggests a single common index, a shared conception between philosophers and economists of things good and bad in life. These would not only consist of exchangeable goods but could also refer to other levels of development and states of consciousness. If developed further, as suggested by Kopp (1993), to clarify *who* should determine which goods and conditions for a good life make it onto that index, this line of thought could indeed result in a more complete economic theory.

6. ALTERNATIVE APPROACHES

In each of the previous three sections we have outlined an issue area that decision makers need to be aware of when devising public policy that is based on economic evaluations such as CBA. That awareness is not equally called for in all policy domains, as policy decisions in some domains are less vulnerable to our criticisms than in others. It remains up to the judgement of the reader to assess the relevance of the three issue areas and possibly, conclude that CBA can be applied unequivocally to help solve a given policy problem. When decisions have to be made in domains such as those referred to in this chapter, however, policy makers are advised to consider

other methodological approaches that bypass the pitfalls identified. To that end we briefly offer below two alternative approaches. They are not fundamentally new evaluation techniques but are best seen as less stringent variants of CBA and should therefore be easily comprehensible.

In Section 3 we saw that not all costs and benefits that enter economic evaluations can be measured in monetary terms, as some valuation techniques rest on contestable assumptions regarding the quantification of economic value. As a possible way out of this impasse, the policy maker could replace CBA with a similar technique, that of cost utility analysis (CUA). The difference is that, while CBA converts benefits into a monetary metric as a common unit, CUA expresses benefits in terms of the utility they provide to the individual—such as QALYs in the case of health care. It is a non-monetary concept for estimating the value to society of improvements in a status of well-being and thus sidesteps the problem of monetary conversion.

Its merits as a non-monetary economic evaluation technique notwithstanding, CUA remains, just as CBA is, vulnerable to the criticisms we raised in Sections 3 and 4: calculating utility ratings by quizzing individuals for their preferences of well-being is contestable because these preferences might be non-authentic, malformed, strategically motivated, or simply uninformed. And individuals differ—across lives and across stages of their own life—in how they value particular states of well-being. Any attempt to aggregate such incommensurable attributes into a single standard brings about methodological as well as ethical issues.

To cater to these objections, cost effectiveness analysis, or CEA recommends itself as yet another evaluation technique. Both CBA/CUA as well as CEA are formal methods for comparing the benefits and costs of a policy program. The difference is that, while CBA and CUA convert these benefits into monetary value and utility respectively as a common unit, CEA expresses benefits as such, i.e. in terms of a natural unit as some standard of outcome. In the case of health care such an outcome could, for example, constitute the incremental reduction in mortality rate or the increase in the number of immunizations delivered, rather than the monetary value or utility that CBA/CUA would calculate for each of these effects. In the case of environmental regulation an outcome could, for example, constitute the level of air quality as measured by the ambient ozone level, rather than the economic value or utility it provides to humanity. CEA thus sidesteps the problem of monetary conversion as found in CBA *and* the problem of preference satisfaction and utility aggregation as found in CUA.[6]

The detour comes at a price, however, because CEA is a much less powerful tool than CBA or CUA. It can only assess alternative policies where costs relate to a single common effect as measured on a natural scale (such as mortality rate) which may differ in magnitude among the policy options evaluated. It can then be used to choose among those options in terms of their effectiveness-to-cost ratio. Conversely, if the budget is predetermined, that is the costs are "fixed," it can again, only be used to compare various policy options as to their rate of attaining that non-quantified

---

[6] Note that some authors and literatures treat CUA as a particular case of CEA, or CEA and CUA as particular cases of CBA. The three techniques may therefore appear under different labels.

goal, such as decreasing mortality. What it cannot do is to give an indication *how much* should be spent to achieve a policy outcome. Neither can CEA give guidance whether a policy intervention is worth doing at all, for it tacitly assumes that the objective has been deemed worth meeting beforehand. It therefore does not specify how far a program's ratio of effects to costs can fall before it is no longer worth doing. To determine whether resources have been allocated in such a way that benefits to society have been maximized is not possible with CEA.

What neither CBA, CUA, nor CEA can solve, however, is the intrinsic value problem that we addressed in Section 5. Intrinsic values are not merely not commensurable, they are more fundamentally, also not comparable with other benefits and costs. All too often, they are therefore "forgotten" in economic evaluations although they should be allowed to restrict the projects that government may permissibly carry out. In policy practice, such side constraints can be feasibly implemented by giving a veto power to the individuals impacted by the proposed policy. It does not follow, of course, that such rights automatically override any possible net benefits of a proposed policy, but neither are they morally irrelevant.

7. CONCLUSION

In concluding, economic tools are very general techniques that have very stringent information requirements not all of which can always be met. They can therefore not function as a fundamental standard of choice among policy options. This is not a reason to reject economic evaluations per se as they do provide us with information that is morally relevant and thus possibly uncovers hitherto concealed judgements by policy makers eager to cater to special interests. It is, we have argued, both unethical and irrational in general to ignore the cost and benefits of a pending policy decision. Yet, it is a reason to acknowledge that economic evaluations should be understood as an input into, rather than a substitute for political deliberation and judgement (Sunstein 2002). Not all situations call on us to maximize value. Some simply compel us to respect it. Economic evaluations should be seen as a useful heuristic to raise red flags about policy proposals and identify the economic factors involved. Whether economic factors are, in fact, the dominant concern at all in a given situation is a judgement that will have to remain within the realm of responsibility of the policy maker.

REFERENCES

ADLER, M. D., and POSNER, E. A. (eds.) 2001. *Cost Benefit Analysis: Legal, Economic, and Philosophical Perspectives.* Chicago: University of Chicago Press.

ARROW, K. J. 1963. Uncertainty and the welfare economics of medical care. *American Economic Review*, 53(5): 941–73.