European Commission

JRC TECHNICAL REPORTS

On the use of Cost-Benefit Analysis and Multi-Criteria Evaluation in *ex-ante* Impact Assessment

Munda, Giuseppe

2017

EUR 28768 EN

Contact information
Giuseppe Munda
European Commission, Joint Research Centre
Directorate I – Unit JRC.I.1 – Modelling, Indicators and Impact Evaluation
TP 361 – Via E .Fermi 2749 – I-21027 – Ispra (Va) – ITALY
giuseppe.munda@ec.europa.eu

JRC Science Hub
https://ec.europa.eu/jrc

JRC107900

EUR 28768 EN

PDF ISBN 978-92-79-73213-3 ISSN 1831-9424 doi:10.2760/311199

Luxembourg: Publications Office of the European Union, 2017

How to cite this report: Munda G. *On the use of Cost-Benefit Analysis and Multi-Criteria Evaluation in ex-ante Impact Assessment,* EUR 28768 EN, Publications Office of the European Union, Luxebourg, 2017, ISBN 978-92-79-73213-3, doi:10.2760/311199, JRC107900

Contents

Acknowledgements

Comments by Leen Hordijk are gratefully acknowledged. The usual disclaimer applies.

Note

This report has been developed in the context of the activities of the Competence Centre on Modelling.

Abstract

When a public administration wishes to implement policies, there is a need of comparing different options and valuating and evaluating them to assess their social attractiveness. Traditionally, welfare economics has used cost-benefit analysis based on the Kaldor-Hicks compensation principle, which was invented to achieve two clear objectives:

1) To compare individuals' preferences according to the efficiency oriented utilitarian calculus, explicitly avoiding the principle one individual, one vote.

2) To implement an objective evaluation criterion, that could be accepted in the framework of the positivistic philosophical paradigm.

A relevant question now is: are these objectives still relevant in the 21st Century? This article aims at presenting a structured comparison of the main distinguishing features of monetary and non-monetary methods useful in the assessment of public policy options. In particular, cost-benefit analysis (CBA) and multi-criteria evaluation (MCE) are discussed according to the following ten comparison criteria: efficiency, fairness, democratic basis, effectiveness, problem structuring, alternatives taken into account, policy consequences, comprehensiveness, transparency and mathematical aggregation rule.

A conclusion is that regarding public policy problems, CBA and MCE can be considered as competitive methods only if all consequences of a policy decision can be correctly transformed into monetary values and efficiency is the only relevant policy objective. In all other cases, CBA can be used as a criterion in a MCE framework; thus in general terms, CBA and MCE are complementary in nature. MCE seems to be an appropriate public policy framework to integrate different scientific languages, when concerns about civil society and future generations have to be considered along with policy objectives and market conditions.

1 Cost-Benefit Analysis: Efficiency, Fairness and Democratic Basis

This report aims at presenting a structured comparison of the main distinguishing features of monetary and non-monetary methods useful in an *ex-ante* impact assessment framework. In particular, cost-benefit analysis (CBA) and multi-criteria evaluation (MCE) are discussed according to ten comparison criteria: efficiency, fairness, democratic basis, effectiveness, problem structuring, alternatives taken into account, policy consequences, comprehensiveness, transparency and mathematical aggregation rule. This Section deals with efficiency, fairness and democratic basis in the framework of cost-benefit analysis. Next Section deals with the same three comparison criteria in the framework of multi-criteria evaluation. Section 3 presents a systematic comparison between CBA and MCE according to the other seven criteria. Finally some conclusions are illustrated.

When the European Commission (or any other public administration) wishes to implement policies, there is a need of comparing different options and valuating and evaluating them to assess their social attractiveness. A monetary valuation is characterised by an attempt to measure all effects in monetary units, whereas a non-monetary evaluation utilises a wide variety of measurement units to assess such effects. Cost-benefit analysis (CBA) and cost-effectiveness analysis (CEA) are well-known examples of monetary valuation methods (Munda *et al*., 1995).

The rationality behind cost-benefit analysis assumes that any individual makes rational decisions only if she/he adds the advantages and disadvantages of a particular action, so that some kind of best decision can always be made. Of course, the essence of cost-benefit analysis is that it is not confined to decisions that affect one individual; it relates to social decisions. Then, does the characteristic of rationality remain if we extend it to the social context? The basic argument underlying CBA is that this rationality does remain. That is, if individuals are left free to carry out their own personal cost-benefit analyses in respect to a given policy, then we can simply aggregate the results to secure a social assessment. It is the emphasis on this social view that generates many of the philosophical problems of CBA. First, what set of individuals constitutes society? More important, only the individuals of the present society are counted. On this argument, CBA could be undemocratic if it is judged on behalf of future generations.

The notion of individual preference which is relevant to the Kaldor-Hicks compensation principle (and cost-benefit analysis), is the preference expressed on the market place (or which would be expressed if there were a market), and not the preference expressed by a political vote (see e.g. Mishan, 1971; Pearce and Nash, 1989). This kind of "economic democracy" is preferred to classical political voting procedures for different reasons:

1. The Kaldor-Hicks principle declares a social state ***A*** "socially preferable" to an existing social state ***B*** if those who gain from the move to ***A*** can compensate those who lose and still have some gains left over. Such a situation is consistent with a Pareto improvement since we have ***B*** indifferent to ***A*** for the losers (once they are compensated) and ***A*** preferred to ***B*** for the winners (if they can over-compensate). If the monetary value of benefits exceeds the monetary value of costs, then the winners can hypothetically compensate the losers and still have some gains left over. The excess of gains over required compensation is equal to the net benefits of the project. While in political voting, minorities always lose since they have to accept "majority dictatorship", economic democracy, implemented through the Kaldor-Hicks compensation principle, always compensate losers, thus it seems to improve the fairness of the policy process.
2. Political systems other than in very well-defined referenda involve voting not for issues so much as for individuals to represent the constituent's view. Market or economic voting is considered closer to the voters' intentions: by definition if a voter*, identified as a consumer*, does not want something, she/he does not buy it.

3. Even if referenda were desirable, they cannot be held continuously on every policy decision that has to be made. To observe consumers' behaviour on the market is much cheaper, quicker and easier.
4. As clearly expressed by Pearce and Nash (1989, p. 7): *"the use of money values permits some expression of the intensity of preference in the vote: it enables the individual to say how deeply he wants or does not want the project or good in question".*

In summary, even though Kaldor and Hicks were looking for a policy criterion implementing objective Pareto efficiency, *explicitly not based on egalitarian considerations,* economic democracy seems to perform much better than political democracy. In fact it is easier and cheaper to implement, it uses more information on individual's preferences (i.e. intensity of preference), and losers are always compensated. Let's then see if these arguments are correct.

The main underlying idea of using preferences expressed on the market is that individuals can be compared by means of a common property, *being consumers*, and one measurement unit i.e. *money values measuring their willingness to pay* for a good or service. One obvious consideration is that the comparison of individuals is possible according to the characteristics of this property and measurement unit only: money values are worth to be used when they are connected to *one objective* and *one institution* only, i.e. economic efficiency and markets. However, they fail to incorporate other objectives and values, such as sustainability or fairness (Munda, 2016).

Choosing any particular operational definition for *value* and its corresponding valuation technique involves making a decision about what is relevant and important. *A priori,* there is no reason why this issue of existence of a plurality of values should be considered a problem that can be solved by considering consumers' preferences as the only relevant social values. Sagoff (1988) made clear the point that one's preferences as a consumer may differ significantly from one's preferences as a citizen.

When one wishes to preserve e.g. a monument or a natural area, a fundamental question is: *is there any resource which society is willing to assign to this objective?* Indeed no society can avoid the economic problem of *"opposition between tastes and obstacles",* as Pareto made clear. To answer this question the concept of total economic value becomes immediately relevant. To attribute monetary values to e.g. historical heritage implies to capture user (actual, option and bequest) and non-user (existential, symbolic, etc.) *values*. However, one should remember that the market alone may be successful in *efficient* allocation of resources, but does not give any guarantee for *preservation* of the cultural or natural heritage at all nor for the *fairness* of the decision taken. According to the compensation principle, the social cost of a given policy option is defined as the sum of money paid as compensation to those who have been suffered damage, the level of utility that the damaged had before the event took place should determine the amount of compensation to pay[1].

[1] Although there are symbolic goods which may present difficult possibilities of transactions in actual or fictitious markets surely. How much one should receive to accept compensation for the destruction of the Big Ben, the Sagrada Familia, the Statue of Liberty or the Coliseum? Indeed Kaldor admitted the existence of such losses of a symbolic kind: *"An increase in the money value of the national income (given prices) is not, however, necessarily a sufficient indication of this condition being fulfilled: for individuals might, as a result of a certain political action, sustain losses of a non-pecuniary kind- e.g., if workers derive satisfaction from their particular kind of work, and are obliged to change their employment, something more than their previous level of money income will be necessary to secure their previous level of enjoyment; and the same applies in cases where individuals feel that the carrying out of the policy involves an interference with their individual freedom. Only if the increase in total income is sufficient to compensate for such losses,*

Negative externalities to be internalised by monetary compensation can also be seen as *"cost-shifting"*. In general, if the damaged people are poor (or even not yet been born), the cost of the internalization of the externality will be low. Indeed the inseparability between efficiency and distributional equity (and thus the unfairness of the decision process) can be easily proved from a formal point of view too. The use of willingness to pay in money terms as a measure of individuals' intensity of preference would be correct only if individuals' income could be measured on a ratio scale of measurement, that is the only degree of freedom is the unit of measurement and not the origin[2]. In this measurement framework if individual ***X*** expresses the twice willingness to pay for the good ***A*** than individual ***Z***, then it is correct to derive that she/he has the double intensity of preference towards that good with respect to ***Z***. Now, we have to consider that although it is true that zero money would be the common origin and thus money could be measured on a ratio scale logically, *on the ontological side,* the real origin of the scale is the true-life individuals' income, which is necessarily different across individuals. Real-world marginal utility of income across individuals is not constant clearly, thus different intensities of preference cannot be compared, on objective grounds, unless we know the exact personal distribution of income.

In a CBA framework, costs and benefits are aggregated linearly in a net present value (NPV) formula. The assumption underlying the NPV rule is that of an additive social welfare function, such as

$$SW = \sum_h U_h \qquad (1)$$

---

and still leaves something over to the rest of the community, can it. be said to be " justified " without resort to interpersonal comparisons" (Kaldor, 1939, p. 551).

[2] The word *measurement* is usually reserved for the situation in which a number is assigned to each observation; this number reflects a magnitude of some quantitative property (how to assign this number constitutes the so-called *representation problem*). The measurement procedure used constitutes a function rule $m: O \to R$, telling how to give an object *o* its *m(o)* value in a systematic way. Measurement operations or procedures differ in the information that the numerical measurements themselves provide about the true magnitudes. Quantitative measurement procedures associate objects $o \in O$ with a real number *m(o)* allowing much more precise statements about the true magnitudes than ordinal scale measurements. Suppose that the statement of equation (I) is true:

$$\begin{cases} m(o_1) \neq m(o_2) \; only \; if \;\; t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \; only \; if \;\; t(o_1) > t(o_2) \\ t(o) = x \; iff \; m(o) = ax + b, where \; a \in R^+ \end{cases} \qquad (I)$$

That is, the numerical measurement *m(o)* is some *affine function* of the true magnitude *x*. When (I) applies, the measurement operation is called *interval scaling*, or measurement at the *interval-scale level.* When measurement is at the interval-scale level, any of the ordinary operations of arithmetic can be applied to the differences between numerical measurements, and the results can be interpreted as statements about *magnitudes* of the underlying property. It is sometimes possible to find measurement operations making the statement of Equation (II) true:

$$\begin{cases} m(o_1) \neq m(o_2) \; only \; if \;\; t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \; only \; if \;\; t(o_1) > t(o_2) \\ t(o) = x \; iff \; m(o) = ax, where \; a \in R^+ \end{cases} \qquad (II)$$

When the measurement operation defines a function such as the statement contained in (II), then measurement is said to be at the *ratio-scale level*. For such scales, ratios of numerical measurements are unique and can be interpreted directly as ratios of magnitudes of objects.

where the subscript *h* denotes the individual to whom the utility function applies. Under the assumption that the marginal utility of money income (λ) is *equal for all individuals,* the variation of this social welfare function indicating the social worth of a project is:

$$\Delta SW = \sum_h \sum_i \frac{\partial U_{ih}}{\partial Y_{ih}} \bullet \Delta Y_{ih} = \lambda \sum_h \sum_i P_i \Delta Y_{ih} = \lambda \sum_i P_i \Delta Y_i \qquad (2)$$

where *h* subscript denotes the individual to whom the utility function and quantity of the good Y_i apply. The translation into monetary terms is accomplished by the equation (3)

$$\lambda \frac{\partial U_i}{\partial Y_i} = P_i \qquad (3)$$

where P_i is the (relative) price of good *i*.

Obviously, the assumption of the constancy of the marginal utility of income across individuals is a distributional question, and that assumption embodies particular social values. Given that society is unlikely to be indifferent among various possible distributions of income, some ways of integrating the distributional aspects into the analysis have to be found. The most popular methodology is to introduce distributional weights explicitly, by using different weights for different social groups (Bojo *et al.,* 1990). However, it is not clear how to derive such weights, since they can be based on a variety of ethical, philosophical and methodological principles and who should attach them (economists, policy-makers, society, ...). On the other hand, one has to note that failures to use any weighting system imply making the *implicit value judgement* that the existing distribution of income is optimal. If, and only if, one is happy with such a value judgement, it is reasonable to use un-weighted market valuations to measure costs and benefits. *Therefore, there is no escape from value judgements,* the compensation principle is not the positivistic objective evaluation criterion Hicks hoped to be. On the other side it does not consider individuals as equal exactly the goal Kaldor aimed at, it can be considered a direct application of the ancient principle that property owners should count more (see Munda, 2017a for more details on this point).

2 Multi-Criteria Evaluation: Efficiency, Fairness and Democratic Basis

The most important non-monetary approach to *ex-ante* Impact Assessment (IA) is no doubt multi-criteria evaluation (MCE). The basic methodological foundation of MCE is *incommensurability,* i.e. the idea that in comparing options, a plurality of dimensions and perspectives are normally needed, and thus there is an irreducible value conflict when deciding what common comparative term should be used to rank alternative actions. For example, when assessing urban well-being of various cities, this is not evaluated as good or bad as such, but rather, as good, bad, beautiful or ugly in relation to different descriptions. It can be at one and the same time a "good average income" and a "bad social inclusion", a "beautiful skyline" and an "ugly cultural heritage".

The basic idea of multi-criteria evaluation (MCE) is to achieve the comparability of incommensurable values. From an operational point of view, the major strength of MCE is its ability to deal with policy issues characterised by various conflicting evaluations, thus allowing for an integrated assessment of the problem at hand. Being a decision tool, MCE focuses on the issue of the *opportunity cost* connected to the choice of any policy option, thus efficiency is surely an important objective to be considered. Differently from economic efficiency assessment tools such as CBA or frontier methods such as Data Envelopment Analysis (DEA), traditionally used in operational research, MCE is based on a multidimensional definition of efficiency, where inputs and outputs are not transformed into a single measurement rod. Of course, CBA or DEA could also be one of the criteria used in a MCE exercise, but never the only ones. A clear advantage of MCE is that different objectives, such as efficiency, equity or sustainability can be treated separately in a transparent way.

A *"discrete multi-criterion problem"* can be formally described as follows (see e.g. Figueira *et al.,* 2016). *A* is a finite set of *N* feasible actions (or alternatives). *M* is the number of different points of view, or evaluation criteria, $g_{m,}$ that are considered relevant to a specific policy problem. Where action ***a*** is evaluated to be better than action ***b*** (both belonging to the set *A*), by the *m*-th point of view, then $g_m(a)>g_m(b)$. In this way a decision problem may be represented in an *N* by *M* matrix *P* called an *evaluation or impact matrix*. In such a matrix, the typical element p_{mn} *(m=1, 2 , ... , M; n=1, 2 , ... , N)* represents the evaluation of the *n*-th alternative by means of the *m*-th criterion, in other words, each criterion score represents the performance of each alternative according to each criterion (see Table 1). The impact matrix may include quantitative, qualitative or both types of information.

		Alternatives			
Criteria	**Units**	$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$	$\mathbf{a_4}$
$\mathbf{g_1}$		$g_1(a_1)$	$g_1(a_2)$	.	$g_1(a_4)$
$\mathbf{g_2}$		.	.	.	.
$\mathbf{g_3}$		.	.	.	.
$\mathbf{g_4}$		.	.	.	.
$\mathbf{g_5}$		.	.	.	.
$\mathbf{g_6}$		$g_6(a_1)$	$g_6(a_2)$	.	$g_6(a_4)$

Table 1. Example of an Impact Matrix

In a discrete multi-criteria problem, there is a range of multi-criteria problem formulations, which may take one of the following forms (Roy, 1996):

(α) the aim is to identify one and only one final alternative;

(β) the aim is the assignment of each alternative to an appropriate predefined category according to what one wants it to become afterwards (for instance, acceptance, rejection or delay for additional information);

(γ) the aim is to rank all feasible alternatives according to a total or partial pre-order;

(δ) the aim is to describe relevant alternatives and their consequences.

In synthesis, the information contained in the impact matrix useful for solving the so-called multi-criterion problem is:

- Intensity of preference (when quantitative criterion scores are present).
- Number of criteria in favour of a given alternative.
- Weight attached to each single criterion.
- Relationship of each single alternative with all the other alternatives.

Combinations of this information generate different aggregation conventions, i.e. manipulation rules of the available information to arrive at a preference structure. As noted by Arrow and Raynaud (1986), in the case where all criteria have ordinal impact scores, if one considers the evaluation criteria as voters, a multi-criteria impact matrix and a voting matrix are identical. As a consequence all results of social choice also apply to multi-criteria evaluation fully; in particular Arrow's impossibility theorem stating that there is no perfect mathematical aggregation rule. Thus, unlike other mathematical fields, only "reasonable" mathematical procedures can be developed in this framework. Reasonable here means that algorithms can be evaluated not only according to the *formal properties* they present, but, overall, according to the *empirical consequences* implied by their use.

Historically the first stage of the development of MCE is characterised by the so-called methodological principle of *multi-criteria decision making (MCDM).* The main aim of this is to elicit clear subjective preferences from a mythical decision-maker, and then try to solve a well-structured mathematical decision problem by means of a, more or less, sophisticated algorithm. In this way a multi-criterion problem can still be presented in the form of a classical optimization problem (Keeney and Raiffa, 1976).

The limitations of the classical concept of an optimum solution and the consequential importance of the *decision process* were emphasised by authors such as Herbert Simon and Bernard Roy. According to Roy (1996) saying that a decision is a good or bad one is in general impossible on the basis of referring only to a mathematical model. All aspects of a decision process which leads to a given decision also contribute to its quality and success. Thus, establishing the validity of a procedure is impossible, either based on a notion of *approximation* (i.e., discovering pre-existing truths) or on a mathematical property of *convergence* (i.e., does the decision automatically lead, in a finite number of steps, to the optimum *a**?). The final solution is more like a "creation" than a discovery. Under the concept of a *Multiple-Criteria Decision Aid (MCDA)* the principal aim is not to discover a solution, but to construct or create something which is viewed as liable to help an actor taking part in a decision process either to shape, argue, and/or transform her/his preferences, or to make a decision in conformity with his/her goals (Roy, 1996).

The need for public participation has been increasingly recognised in MCE. In particular, *Social Multi-Criteria Evaluation (SMCE)* recognises the need to extend MCDA by incorporating the notion of the social actor. Thus, a SMCE process must be as participative and as transparent as possible; although, participation is a necessary but not a sufficient condition for successful evaluation (Munda, 2004; 2008). This is the main reason why the concept of SMCE is proposed in place of Participatory Multi-Criteria Evaluation or Stakeholder Multi-Criteria Decision Aid (Banville *et al.,* 1998). The strength of SMCE is the fact that the use of various evaluation criteria has a direct translation in terms of plurality of values used in the evaluation exercise.

3 A Systematic Comparison Between Cost-Benefit Analysis and Multi-Criteria Evaluation According to Other Seven Criteria

In this section, a comparison of the key characteristics of cost-benefit analysis and multi-criteria evaluation will be carried out on the base of other seven comparison criteria, that is: effectiveness, problem structuring, alternatives taken into account, policy consequences, comprehensiveness, transparency and mathematical aggregation rule.

a) Effectiveness

Commission Impact Assessment correctly considers both the objectives of effectiveness and efficiency; in fact it is of key importance to understand that efficiency alone cannot be a relevant policy objective. *Effectiveness* (i.e. the degree to which goals and levels of output are achieved or problems are solved) is at least equally important; otherwise there is the risk to drive the policy evaluation framework towards a situation where efficiency means just "cheap" (Agasisti *et al.,* 2017). For this reason, it is important to have a clear understanding of the difference between efficiency and effectiveness. To clarify this point, let's consider the following four situations obtained by combining efficiency with effectiveness in a public policy framework:

Effective	**A)** Goals are achieved, (e.g. high education levels, good environmental quality standards, low percentage of population at risk of social exclusion…) but too many resources are used. The system is effective but there is a waste of resources.	**B)** Goals are achieved by using a reasonable amount of resources. Of course, this is the best situation. An obvious issue is the definition of what a "reasonable amount" means.
Non-effective	**C)** Goals are not achieved and a lot of resources are used. This is the worst situation.	**D)** Goals are not achieved but the amount of resources used is low. The system looks efficient (because it uses wisely poor resources) but it is non-effective. In this case efficiency is confused with parsimony.
	Non-Efficient	**Efficient**

It is immediately evident that efficiency is a relevant policy objective only and only if it is considered in combination with effectiveness; otherwise two different systems (e.g. countries, regions, cities,…) might present the same level of efficiency, with very different levels in goal achievements.

Public policies based on monetary principles sometimes might be operative, but one should be very cautious in applying such principles as a general guideline For example, if the policy objective is to reduce the tourist pressure on Venice, one may think of limiting

the number of visitors by imposing the payment of an entry ticket and to use the money collected to maintain the city's cultural heritage. However, one could argue that due to the *relative scarcity* of a peculiar economic good as Venice, people will be willing to pay the price of the ticket anyway. Thus, the economic instrument *entry ticket* will be useful for collecting money, but not for reducing the tourist pressure. Market based policy instruments alone may be successful in efficient allocation of resources, but do not give any guarantee for preservation of the cultural or natural heritage at all. Once something is on the market, it can be bought or sold and so the willingness to accept and the compensation principle may easily cause the destruction of any asset.

Monetary compensation is no doubt the only possible tool when an irreparable and irreversible damage has already occurred. This way, if e.g. an accident with serious contamination occurs (e.g. in the case of Seveso in Italy (1976), of Bhopal in India (1984), of the Exxon Valdez in Alaska (1989), of the oil-tanker Prestige offshore the coasts of Galicia (2002), or the BP oil spill in the Gulf of Mexico (2010)) it seems correct and opportune to compensate the victims of such contamination. But it stays to verify if, in the long run, compensation is an effective tool to prevent the appearance of enormous future social costs. Society has a much longer life expectancy than individuals, thus the value society attaches to e.g. natural resources is likely to deviate from individual values, since the simple summation of individual preferences may imply the extinction of species and ecosystems. This implies that public policy cannot be merely based upon the aggregation of individual values, and estimation of willingness to pay at any particular point of time. It is interesting to note that, Walras himself already noted that the market cannot be used as a basis for *rational collective decision-making* and that "*human destinies are not absolutely independent, but to some extent dependent on one another. There is a social morality which is distinct from individual morality*" (cited in Burgenmeier, 1994, p. 347).

One could argue that, the presence of irreversibility and uncertainty urges us to abandon the compensation principle in favour of the *precautionary principle* (it is more prudent a social conservationist attitude). The application of the precautionary principle introduces some elevated costs surely, but how much would the non-application cost? The burden could be enormous, as stated by the European Environment Agency[3]. There is no doubt that from the viewpoint of society, the application of the precaution principle is ecologically and economically more convenient than a long series of disastrous accidents. Of course, this principle implies that the majority of the society (mainly the non-experts), outside the economic system (i.e. outside the market mechanisms), would decide the *amount* of e.g. cultural or natural capital to be protected. Thus, in the Venice example, the *maximum number of visitors* allowed per day should be clarified, and this can only be done on heuristic grounds since tourist carrying capacity can hardly be computed precisely.

In this context, from an economic point of view, the only instrument left is cost-effectiveness (CEA); that is given a certain *physical* target (e.g. the amount of cultural heritage to be preserved or the amount of contamination to be accepted), it is rational to try to get it by means of the lowest possible use of resources (i.e. at the minimum social cost). Obviously there are several targets possible. In general two rankings are possible:

i. According to the lowest cost.
ii. *According to the physical target* (e.g., the more monuments preserved, the better).

Perhaps a discussion would lead to the judgment that the improvement of a physical target to a better one is worth the extra economic cost, or perhaps the opposite judgment will be reached. In both cases we would have an ordinal ranking of alternatives and CEA would "fall down" into *multi-criteria evaluation,* i.e. two criteria and two different

[3] Late lessons from early warnings: the precautionary principle 1896-2000, European Environment Agency, Environmental issue report, No.22, 2001.

rankings must be explicitly dealt with. In summary, taking effectiveness into account necessarily implies the use of MCE, in no way CBA can deal with it properly.

b) Problem Structuring

The main steps to develop a CBA are:

i. identification of costs and benefits and their monetisation

ii. choice of a social rate of discount

iii. choice of a time horizon

iv. construction of a one-dimensional indicator bringing together all the benefits and costs.

In CBA the main effort consists in trying to apply the right valuation techniques to transform everything into money terms. It has to be remembered that the fundamental question is: *is there any resource, which society is willing to assign to this objective?* To answer this question the concept of *"total economic value"* becomes immediately relevant. To attribute monetary values to e.g. historical heritage implies to capture *user* (actual, option and bequest) and *non-user* (existential, symbolic, etc.) *values*.

Different valuation techniques exist, and it is not always clear which is the best technique to apply in a certain real-world problem; the principal ones are contingent valuation, the travel cost method, hedonic pricing, and the shadow project approach. Among these only *contingent valuation* is universally applicable. The aim of contingent valuation is to elicit valuations (or "bids") which are close to those that would be revealed if an actual market existed. Respondents say that they would be willing to pay or willing to accept if a market existed for the good in question. In order to determine the value of intangible goods and services, economists try to identify how much people would be willing to pay (willingness to pay (WTP)) for these goods in artificial markets. Alternatively, the respondents could be asked to express their willingness to accept (WTA) compensation.

In operational terms, the application of a SMCE framework involves the following main steps:

i. Description of the relevant social actors. For example, institutional analysis may be performed on historical, legislative and administrative documents to provide a map of the relevant social actors.

ii. Definition of social actors' values, desires and preferences. In a SMCE framework, the pitfalls of the technocratic approach can be overcome by applying different methods of sociological research.

iii. Generation of policy options and selection of evaluation criteria as a process of co-creation resulting from a dialogue between analysts and social actors. In this way, evaluation criteria become a technical translation of social actors' needs, preferences and desires.

iv. Construction of the multi-criteria impact matrix synthesising the scores of all criteria for all policy alternatives, i.e. the performance of each option according to each criterion.

v. Construction of an equity impact matrix, including all the distributional consequences of each single option on the various social actors.

vi. Application of a mathematical procedure in order to aggregate criterion scores and obtain a final ranking of the available alternatives.

vii. Finally, sensitivity and robustness analysis looks at the sensitivity of results to the exclusion/inclusion of different criteria, criterion weights and dimensions (see e.g.

Saltelli *et al.,* 2008). While such analysis may look very technical, in reality a social component is always present too. That is, inclusion/exclusion of a given dimension, or set of criteria, normally involves a long story of social, political and scientific controversy, and involves social values and social actors.

These steps are not rigid; problem structuring may vary a lot across different real-world problems. This means that when an attempt is made to model an empirical situation, the presence of a certain subjective component appears to be an inevitable phenomenon. In general, this is a desirable feature, in fact when a model without any creative, personal or subjective influence of a model designer is used, this is inevitably characterised by a certain rigidity which prevents it adhering completely to the situation modelled. This could make it necessary to "force reality" because in the end the tendency will be to make reality fit the model.

On the other hand, to compute total economic values has nothing to do with the idea of a *"true"* or *"correct"* objective value. All monetary valuation attempts suffer deep philosophical problems (see e.g. Copp, 1987; Fusco, 1986; Hansson, 2007; O'Neill, 1993; Sagoff, 1988; Spash, 2008) and technical uncertainties (see e.g. Aldred, 2009; Frey, 1986; Grüne-Yanoff, 2009; Hansen, 2011; Martinez-Alier *et al.,* 1998; Munda, 1996; Vatn and Bromley, 1994) such as:

- Which monetary valuation technique has to be used?
- Which time horizon has to be considered?
- Which social discount rate?

Indeed, the use of models with characteristics of subjectivity or of subjectivism, depends in the latter analysis on the ability and ethical behaviour of the researchers constructing the model. It is important to remember this above all, when MCE or CBA methods are used to "justify" or "defend" political decisions (Funtowicz and Ravetz, 1991).

c) Alternatives Taken Into Account

CBA is limited to discrete problems. Even only one alternative can be evaluated versus a base alternative. Arrow's axiom of independence of irrelevant alternatives[4] is always respected since the degree of attractiveness of each single alternative is independent from all other actions. In MCE, any finite or even infinite number of alternatives can be taken into account. With respect to the axiom of independence of irrelevant alternatives, some methods (e.g. multi-attribute utility theory (MAUT)) respect it; while the results provided by other methods (e.g. outranking methods) are also a function of the set of alternatives considered.

d) Policy Consequences

In CBA it is necessary to identify all costs and benefits and then their correct transformation into monetary values. In a multi-criterion problem, a reasonably large number of criteria, reflecting completely different points of view, can be taken into account. The criterion scores can be quantitative or qualitative. The possibility of taken into account qualitative consequences of decisions is very important in the framework of public policy problems where, intangibles are often present.

e) Comprehensiveness

A variety of policy objectives such as efficiency, equity, effectiveness, sustainability and so on can hardly be incorporated in a cost-benefit analysis simultaneously. MCE is by definition multidimensional in its nature and thus it can easily tackle different and conflicting policy objectives measured on different metrics.

f) Transparency

[4] Arrow's axiom of "the independence of irrelevant alternatives" states that the choice made in a given set of alternatives A depends only on the ordering made with respect to the alternatives in that set. Alternatives outside A (irrelevant since the choice must be made within A) should not affect the choice inside A.

Since in CBA all dimensions are translated into money terms and then often aggregated by using a linear mathematical rule, which is completely compensatory, the possibility of a complete understanding of all the profiles taken into consideration is quite low. However, the possibility of presenting the background data can increase the transparency of a given application. In MCE, all the multidimensional profiles of the problem are clearly shown in the original scales of measurement. From an operational point of view, the great degree of transparency allowed in multi-criteria models is one of its most important factors of success.

g) Mathematical Aggregation Rule

The mathematical axiomatization of cost-benefit analysis is complete, since in a monocriterion analysis it is possible to discover a precise optimal solution, if it exists. Furthermore, CBA is based on standard investment criteria such as Net Present Value (NPV) or the Internal Rate of Return (IRR). However, it is not always clear which investment criterion should be used. If the NPV is used, since it can be considered an additive utility function, the condition of preference independence should always hold[5] (Munda, 1996). Since a multi-criterion problem is by definition mathematically ill-structured, i.e. it has no objective solution, a complete mathematical axiomatization of MCE is very difficult. This is also the most important cause of the flourishing of a lot of different theories and models. This is a weak point of the multi-criterion approach. However, a way of bypassing this problem is to indicate clearly the axiomatic system underlying any method and to list the set of properties considered desirable.

An issue, that makes multi-criterion aggregation conventions intrinsically complex, is the fact they are *formal, descriptive and normative* models simultaneously. As a consequence, the properties of an approach have to be evaluated at least in the light of these three dimensions. Musgrave (1981) in the framework of the debate on the maximisation assumption in microeconomics, made a very useful classification of the assumptions used in economic theory. He makes a distinction among *negligibility assumptions, domain assumptions and heuristic assumptions.* The first type is required to simplify and focus on the essence of the phenomena studied. The second type of assumptions is needed when applying a theory to specify the domain of applicability. The third type is needed either when a theory cannot be directly tested or when the essential assumptions give rise to such a complex model that successive approximation is required. One might see this last type of assumptions as the sake of learning about limits to the relationship between understandable implications and complexity.

Here, I will indicate some properties that can be considered desirable for a discrete multi-criteria aggregation rule (often called multi-criteria *method*) in the framework of Commission *ex-ante* IA (see Munda, 2017b for more details on this point). Of course in another framework, e.g. stock exchange investments, these properties can easily be irrelevant or even undesirable. These desirable properties are:

<u>Descriptive domain assumptions:</u>

- Mixed information on criterion scores should be tackled in the form of ordinal, crisp, stochastic and fuzzy criterion scores.

<u>Normative domain assumptions:</u>

- Simplicity is desirable to guarantee coherence between the problem structuring and the results obtained, and means the use of as less *ad hoc* parameters as possible.
- The most useful result for policy-making is a complete ranking of alternatives.

[5] This property is a necessary condition for the existence of a linear aggregation rule. From an operational point of view this means that an additive aggregation function permits the assessment of the marginal contribution of each cost and benefit separately. Each marginal contribution can then be added together to yield a total value. This implies that among the different aspects of a policy option there are no phenomena of synergy or conflict, this is rather unrealistic from a scientific point of view.

- Weights are meaningful only as importance coefficients and not as trade-offs.
- Complete compensability[6] is not desirable.

Heuristic descriptive assumptions:

- When not all intensities of preference are meaningful, indifference and preference thresholds are useful exogenous parameters[7].
- Dominated alternatives have to be considered[8].

Finally one should note that these selection properties can be applied only to methods who achieve a set of minimum formal requirements, the main important being the following[9].

Formal domain assumptions:

- Unanimity.
- Monotonicity.
- Neutrality.

Negligibility formal assumptions:

- Anonymity.

We may conclude that in a IA framework, Condorcet consistent mathematical rules seem adequate for finding rankings of policy options. They present a low probability of rank reversal and are not allowing for a complete compensability thus weights can be treated as importance coefficients. A weak point is the high probability of presence of cycles; their solution normally implies ad hoc rules of thumb (see Munda, 2008 for a comprehensive analysis of this issue).

[6] Complete compensability is not desirable for the problem we are dealing with, since it implies that e.g. a good performance on efficiency would offset a very bad one on effectiveness or vice versa. It has to be noted that CBA always allows the highest degree of compensability since it is explicitly based on the Kaldor-Hicks compensation principle and costs and benefits are aggregated linearly.

[7] This relates to the famous bald paradox in Greek philosophy (how many hairs one has to cut off to transform a person with hairs to a bald one?), later on Poincaré (1935, p. 69) and finally Luce (1956) made the point that the transitivity of indifference relation is incompatible with the existence of a sensibility threshold below which an agent either does not sense the difference between two objects, or refuses to declare a preference for one or the other. Luce was the first one to discuss this issue formally in the framework of preference modelling. Mathematical characterizations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

[8] This of course applies to discrete methods only and implies that the aggregation rules belong to the family of non-frontier methods.

[9] In social choice, the reaction to Arrow's theorem has been the search for less ambitious voting structures; there is a need to keep a few basic requirements only. These basic requirements are generally three:

1. Anonymity: all criteria must be treated equally.
2. Neutrality: all alternatives must be treated equally.
3. Monotonicity: more support for an alternative cannot jeopardize its success.

One should note that, while anonymity is clearly essential in the case of voters, it is not so in the multi-criterion problem since criterion weights can be normally introduced.

4 Conclusions

Monetary valuation methods e.g. CBA or CEA are based on economic concepts such as consumer's surpluses, market failures, demand curves which are just a partial point of view, since connected with one institution only: *markets*. From a social point of view, issues connected with actions outside of markets and behaviour of people different from the class of consumers should also be taken into account. *The European Commission White Paper on Governance*[10] (where principles such as *transparency*, *participation* and *accountability* are emphasized) goes in this direction. It has to be reiterated that the point is not to be against giving economic value to natural resources, to cultural heritage or even to human life[11]. A location may be valuable for its biodiversity (measured in richness of species or genetic variety), and also as a landscape, and have also economic value (measured by differential rent, and also by the travel cost method, or contingent valuation). These are different types of value. The point is that it is misleading to take social decisions based on only one type of value.

One should note that regarding public policy problems, CBA and MCE can be considered as competitive methods only if all consequences of a policy decision can be correctly transformed into monetary values; but this is very difficult. Thus we can say that, given the presence of unpriced impacts, often MCE is the only possible approach. However when monetary values are present, CBA can be used as a criterion in a MCE framework dealing with the objective of efficiency in a proper way. It is thus possible to conclude that CBA and MCE are complementary in nature (MCE being the most comprehensive one).

In CBA only utility based models of a complete type are used, as a consequence, complete compensability and preference independence are always assumed. In MCE, various mathematical aggregation rules exist; models with completely different properties can be used. This makes multi-criteria evaluation more flexible but also more confusing since a method has to be chosen and the final results may be very sensitive to this step. On the other hand this problem of "method uncertainty" is also present in CBA, since different valuation techniques and aggregation rules (e.g. NPV or IRR) exist.

[10] Communication from the Commission of 25 July 2001 "European governance - A white paper" [COM(2001) 428 final - Official Journal C 287 of 12.10.2001]
http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=LEGISSUM:l10109

[11] One has to note that the issue is not maintaining that a human life has infinite value; for example, a reduction in road accidents can be secured at some cost, but society is unlikely to devote the whole of the national income to this end. The point is that often this valuation is made *implicitly* and stating that is a technical issue, when it is a socio-political one instead.

References

Agasisti, T., R. Hippe and G. Munda (2017) Efficiency of investment in compulsory education: empirical analyses in Europe; EUR 28607 EN. Luxembourg (Luxembourg): Publications Office of the European Union; JRC106678; doi 10.2760/975369.

Aldred J. (2009) Ethics and Climate Change Cost-Benefit Analysis: Stern and After, New Political Economy, Volume 14, Issue 4, pages 469-488.

Arrow, K.J. (1963) Social choice and individual values. 2d edition, Wiley, New York.

Arrow, K.J., Raynaud, H. (1986) Social choice and multicriterion decision making. M.I.T. Press, Cambridge.

Banville C., Landry M., Martel J.M., Boulaire C. (1998) A stakeholder approach to MCDA, Systems Research and Behavioral Science, Vol. 15, pp. 15-32.

Bojo J., Maler K.G. and Unemo L. (1990) Environment and development: an economic approach, Kluwer, Dordrecht.

Burgenmeier B. (1994) The misperception of Walras, American Economic Review, (84)1, pp. 342-352.

Copp D. (1987) The justice and rationale of cost-benefit analysis, Theory and Decision, Volume 23, Issue 1, pp 65-87.

Figueira, J., Greco, S. and Ehrgott, M. (eds.) (2016) Multiple-criteria decision analysis. State of the art surveys. Springer International Series in Operations Research and Management Science, New York.

Frey B.S. (1986) Economists Favour the Price System Who Else Does?, Kyklos, Volume 39, Issue 4, Pages: 537–563

Funtowicz S.O., Ravetz J. R. (1991) A new scientific methodology for global environmental issues, in R. Costanza (ed.) Ecological Economics, New York, Columbia, pp. 137-152.

Fusco Girard, L. (1986) The complex social value of the architectural heritage. Icomos Information, 1, pp. 19-22.

Grüne-Yanoff T. (2009) Mismeasuring the value of statistical life, Journal of Economic Methodology, Volume 16, Issue 2, pages 109-123.

Hammitt J.K. (2013) Positive versus normative justifications for benefit-cost analysis: implications for interpretation and policy, Review of Environmental Economics and Policy, Vol. 7, Issue 2, pp. 199-218.

Hansen F. (2011) The Stern Review and its critics: economics at work in an interdisciplinary setting, Journal of Economic Methodology, Volume 18, Issue 3, pages 255-270.

Hansson S.O. (2007) Philosophical problems in Cost-Benefit Analysis, Economics and Philosophy, 23, pp. 163-183.

Hicks, J.R. (1939) The foundations of welfare economics. Economic Journal, vol. 49, No. 196, pp. 696–712.

Kaldor, N. (1939) Welfare comparison of economics and interpersonal comparisons of utility. Economic Journal, vol. 49, No. 195, pp. 549–552.

Keeney R., Raiffa H. (1976) Decision with multiple objectives: preferences and value trade-offs, Wiley, New York.

Luce R.D. (1956) Semiorders and a theory of utility discrimination, Econometrica, 24, pp.178-191.

Martinez-Alier, J., Munda, G., O'Neill, J. (1998) Weak comparability of values as a foundation for ecological economics. Ecological Economics, 26, pp. 277-286.

Mishan E.J. (1971) Cost-benefit analysis, Allen and Unwin, London.

Munda G., Nijkamp P. and Rietveld P. (1995) Monetary and non-monetary evaluation methods in sustainable development planning, Economie Appliquée, tome XLVIII, N. 2, pp.145-162.

Munda G. (1996) Cost-benefit analysis in integrated environmental assessment: some methodological issues, Ecological Economics, vol. 19, N. 2, pp. 157-168.

Munda G. (2004) "Social multi-criteria evaluation (SMCE)": methodological foundations and operational consequences, European Journal of Operational Research Vol. 158, Issue 3: 662- 677.

Munda G. (2008) Social multi-criteria evaluation for a sustainable economy, Springer, Heidelberg, New York.

Munda G. (2016) Beyond Welfare Economics: Some Methodological Issues, Journal of Economic Methodology, Volume 23, Issue 2, Pages 185-202.

Munda G. (2017a) Dealing with fairness in public policy analysis, EUR 28751 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/75185.

Munda G. (2017b) A social multi-criteria framework for ex-ante impact assessment, EUR 28752 EN, Publications Office of the European Union, Luxembourg, doi:10.2760/909528.

Musgrave A. (1981) Unreal assumptions in economic theory: the F-twist untwisted, Kyklos, 34, pp. 377-387.

O'Neill, J. (1993) Ecology, Policy and Politics. Routledge, London.

Pearce D.W., Nash C.A. (1989) The social appraisal of projects, MacMillan, London.

Poincaré H. (1935) La valeur de la science, Flammarion.

Roubens M., Vincke Ph. (1985) Preference modelling, Springer-Verlag, Heidelberg.

Roy, B. (1996) Multicriteria methodology for decision analysis. Kluwer, Dordrecht.

Sagoff M. (1988) Some problems with environmental economics, Environmental Ethics, 10, pp. 55-74.

Saltelli, A., Ratto, M., Andres, T., Campolongo, F., Cariboni, J., Gatelli, D., Saisana, M., Tarantola, S. (2008) Global Sensitivity Analysis. The Primer, England: John Wiley & Sons.

Spash, C.L. (2008) Deliberative monetary valuation and the evidence for a new value theory, Land Economics, 84 (3), pp. 469-488.

Vatn A., Bromley D.W. (1994) Choices without prices without apologies, Journal of Environmental Economics and Management, 26, pp. 129-148.

List of abbreviations and definitions

CBA	Cost-Benefit Analysis
CEA	Cost-Effectiveness Analysis
DEA	Data Envelopment Analysis
IA	Impact Assessment
IRR	Internal Rate of Return
MAUT	Multi Attribute Utility Theory
MCE	Multi-Criteria Evaluation
NPV	Net Present Value
SMCE	Social Multi-Criteria Evaluation
WTP	Willingness To Pay
WTA	Willingness To Accept

List of tables

Europe Direct is a service to help you find answers to your questions about the European Union.

Freephone number (*)**:**

00 800 6 7 8 9 10 11

(*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

More information on the European Union is available on the internet (http://europa.eu).

HOW TO OBTAIN EU PUBLICATIONS

Free publications:

- one copy:
  via EU Bookshop (http://bookshop.europa.eu);
- more than one copy or posters/maps:
  from the European Union's representations (http://ec.europa.eu/represent_en.htm);
  from the delegations in non-EU countries (http://eeas.europa.eu/delegations/index_en.htm);
  by contacting the Europe Direct service (http://europa.eu/europedirect/index_en.htm) or
  calling 00 800 6 7 8 9 10 11 (freephone number from anywhere in the EU) (*).

  (*) The information given is free, as are most calls (though some operators, phone boxes or hotels may charge you).

Priced publications:

- via EU Bookshop (http://bookshop.europa.eu).

KJ-NA-28768-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/311199

ISBN 978-92-79-73213-3