Unitat d'Història Econòmica
UHE Working Paper 2013_01

On the Foundations of Well-Being Economics and Policy

Giuseppe Munda

Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, Edifici B, 08193, Bellaterra (Cerdanyola), Spain

E-mail: Giuseppe.Munda@uab.cat

10/04/2013

Edifici B – Campus de Bellaterra
08193 Cerdanyola del Vallès, Barcelona, Spain
Tel.:(+34) 935811203; Fax: (+34) 935812012
http://www.h-economica.uab.es

UAB Universitat Autònoma de Barcelona

Universitat Autònoma de Barcelona

Giuseppe Munda, 2013
On the Foundations of Well-Being Economics and Policy
UHE Working Paper 2013_01
http://www.h-economica.uab.es/wps/2013_01.pdf

Unitat d'Història Econòmica
Departament d'Economia i Història Econòmica
Edifici B, Campus UAB
08193 Cerdanyola del Vallès, Spain
Tel: (+34) 935811203
http://www.h-economica.uab.es

Universitat Autònoma de Barcelona

On the Foundations of Well-Being Economics and Policy

Giuseppe Munda

Departament d'Economia i d'Història Econòmica, Universitat Autònoma de Barcelona, Edifici B, 08193, Bellaterra (Cerdanyola), Spain

E-mail: Giuseppe.Munda@uab.cat

ABSTRACT: When one wishes to implement public policies, there is a previous need of comparing different actions and valuating and evaluating them to assess their social attractiveness. Recently the concept of well-being has been proposed as a multidimensional proxy for measuring societal prosperity and progress; a key research topic is then on how we can measure and evaluate this plurality of dimensions for policy decisions. This paper defends the thesis articulated in the following points:

1. Different metrics are linked to different objectives and values. To use only one measurement unit (on the grounds of the so-called commensurability principle) for incorporating a plurality of dimensions, objectives and values, implies reductionism necessarily.
2. Point 1) can be proven as a matter of formal logic by drawing on the work of Geach about moral philosophy. This theoretical demonstration is an original contribution of this article. Here the distinction between predicative and attributive adjectives is formalised and definitions are provided. Predicative adjectives are further distinguished into absolute and relative ones. The new concepts of set commensurability and rod commensurability are introduced too.
3. The existence of a plurality of social actors, with interest in the policy being assessed, causes that social decisions involve multiple types of values, of which economic efficiency is only one. Therefore it is misleading to make social decisions based only on that one value.
4. Weak comparability of values, which is grounded on incommensurability, is proved to be the main methodological foundation of policy evaluation in the framework of well-being economics. Incommensurability does not imply incomparability; on the contrary incommensurability is the only rational way to compare societal options under a plurality of policy objectives.
5. Weak comparability can be implemented by using multi-criteria evaluation, which is a formal framework for applied consequentialism under incommensurability. Social Multi-Criteria Evaluation, in particular, allows considering both technical and social incommensurabilities simultaneously.

Keywords: Value Theory, Incommensurability, Public Policy, Sustainability, Cost-Benefit Analysis, Multiple Criteria Evaluation, Complexity Theory

JEL Classification: A13, C44, D46, E01, H43, Q58

"It was the best of times, it was the worst of times,
it was the age of wisdom, it was the age of foolishness,
it was the epoch of belief, it was the epoch of incredulity,
it was the season of Light, it was the season of Darkness,
it was the spring of hope, it was the winter of despair,
we had everything before us, we had nothing before us,
we were all going direct to Heaven, we were all going direct
to the other way – in short, the period was so far like the present period,
that some of its noisiest authorities insisted
on its being received, for good or for evil, in the
superlative degree of comparison only".

Charles Dickens – *A Tale of Two Cities*, Signet Classic, New American Library, New York, 1960, p. 13.

1. Introduction

Economic development implies the creation of new assets in terms of physical, social and economic structures. Within a process of "creative destruction" traditional environmental, social, and cultural assets derived from a society's common heritage may disappear. The existence of a plurality of social actors, with interest in the policy being assessed, generates a conflictual situation. "... Looking at a single individual, ... He is prevented from being better off than he is, not only because total production is limited, but also because so much of total production is at the disposal of persons other than himself. The same thing holds, of course, for any group or society of individuals, so long as that group is less than the totality of a closed community" (Hicks, 1939, p. 698-699). An obvious question to start with here is then: how can we evaluate the societal attractiveness of different policy options?

When one wishes to implement public policies, there is a previous need of comparing different actions and valuating and evaluating them to assess their social attractiveness. One of the key tasks of welfare economics is exactly this valuation and evaluation exercise. To define what exactly valuation and evaluation connote is not an easy task, " … we value when comparing objects and evaluate when comparing the relative merits of actions. ... There is a sense in which valuation is passive, while evaluation signifies more of an active engagement. We frequently value in order to evaluate. But not always. We sometimes value simply because we wish to understand a state of affairs, such as the quality of life in a country. Welfare economics studies life's quality, valuing objects and evaluating policies being only a means to measuring the quality of life and to discovering ways to improve it" (Dasgupta, 2001, p. C1).

Traditional welfare economics proposes the measurement of social costs and benefits made on the basis of the so called compensation principle (usually associated with the names of Kaldor (1939) and Hicks (1939). According to this principle, the social cost of a given event 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. Although there are symbolic goods without any possibility of transactions in actual or fictitious markets surely. Who would be willing 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 non-pecuniary kind: "... 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" (Kaldor, 1939, p. 551).

Hicks made very clear the point that *economic welfare* and *social welfare* are very different concepts. In fact he considered a theoretical weakness "… when the reader is asked to accept a direct correlation between economic welfare and social welfare in general (whatever that may be). This is not easy to swallow; in any case it is open to the positivist objection that it reflects a particular social outlook, held by certain classes at certain times, and never likely to be acceptable universally" (Hicks, 1939, p. 697). In recent years, such a concept of social welfare, which Hicks did not appreciate so much for its subjectivity, has gained increasing popularity. A growing quantity of literature has been written about concepts such as *multidimensional poverty* (Sen, 1979, 1985; Duclos *et al.,* 2006), *quality of life, happiness* and *well-being* (e.g. Easterlin, 1995, 2001; Frey and Stutzer, 2002; Michalos, 1980, 1997). This tendency has even increased after the influential Stiglitz, Sen and Fitoussi (2009) report, which proposed the use of the concept of well-being as a proxy for measuring societal prosperity and progress.

A relevant question is hence the following: what are the perspectives to build, on the existing knowledge and consensus, alternative measures of public policy attractiveness which account for the concept of well-being? A starting point may be to connect well-being research agenda with the sustainability one. This allows us to draw upon results already established in the literature and widely accepted by the political and scientific communities; as stated by Arrow *et al.,* (2012), "… Much of the literature on sustainable development has taken human wellbeing to be the object to be sustained." One shared non-controversial result of the sustainability literature is that sustainability is a multidimensional concept, which should at least include economic, social, environmental and institutional dimensions.

The multidimensional nature of well-being is also taken for granted by Stiglitz, Sen and Fitoussi (2009, pp. 14, 15) "To define what well-being means a multidimensional definition has to be used. Based on academic research and a number of concrete initiatives developed around the world, the Commission has identified the following key dimensions that should be taken into account. At least in principle, these dimensions should be considered simultaneously:
i. Material living standards (income, consumption and wealth);
ii. Health;
iii. Education;
iv. Personal activities including work
v. Political voice and governance;
vi. Social connections and relationships;
vii. Environment (present and future conditions);
viii. Insecurity, of an economic as well as a physical nature."

The next point to deal with is whether there is a multidimensional evaluation framework able to cope with all these issues simultaneously. I think that in the economic literature there is some confusion on this issue. In this article I will develop theoretical arguments to prove the desirability of such a multidimensional framework and I will show its existence. In the next Section, I will introduce the concept of commensurability and I will prove that its applicability is very limited. Section 3 will introduce the incommensurability principle as a methodological foundation and an analytic justification of Social Multi-Criteria Evaluation, which is considered as a possible practical framework for well-being policy analysis and in Section 4, some conclusions will be drawn.

2. The commensurability principle as the methodological foundation of comparability

Traditionally, from a philosophical perspective, it is possible to distinguish between the concepts of ***strong comparability*** (there exists a single comparative term by which all different actions can be ranked) implying ***strong commensurability*** (a common measure of the various consequences of an action based on a interval or ratio scale of measurement, such as money or energy) or ***weak***

commensurability (a common measure based on an ordinal scale of measurement, such as consumer's utility), and ***weak comparability***, which implies ***incommensurability*** i.e. there is an irreducible value conflict when deciding what common comparative term should be used to rank alternative options; this irreducible value conflict is unavoidable but compatible with rational choice employing, for example, multi-criteria evaluation (Chang, 1997; Martinez-Alier *et al.,* 1998; Munda, 2004; O'Neill, 1993; Rabinowicz, 2012; Raz, 1986).

In this article, I introduce the new concepts of set and rod commensurability. Let us consider the basic example of apples and oranges, we all learn at primary school. Normally we are thought that we cannot sum up them unless we find a common unit of measurement, i.e. their price or the fact that they both belong to the set of fruits. Commensurability, a necessary condition for strong comparability, can then be implemented in two different ways:

1. By looking for a more general category (set) that can contain *all* the characteristics of the objects we wish to compare; these characteristics are described by using adjectives. I define this as **"*set commensurability*"** *(*e.g. apples and oranges are legitimately lumped together as fruit, along with grapes, bananas, etc.).
2. By finding *one* property common to all objects to be compared and measurable by using one measurement unit; obviously comparison of objects is possible according to the characteristics of this property only. I define this as **"*rod commensurability*"** which can be divided into ***"factual"*** (e.g. since different types of fruits contain sugar, and this can be significant when their juice is extracted for drinking as a liquid; and for this purpose they can be compared with sugarcane, sugar beet, etc.) and ***"potential"*** (if one desires to build an ecologically corrected GDP, there is the need to include non-market goods and services, thus their price has, to some extent, to be invented) ones.

Of course, when possible, set commensurability is the most attractive one since apparently no information is lost in the comparison process, while rod commensurability always requires a kind of reductionism. Let us then start by examining the first approach. Let us start by considering the famous Aristotle's syllogism: *All human beings are mortal, all philosophers are human beings, Socrates is a philosopher thus Socrates is mortal.* By using formal logic, the same syllogism can be written as follows:

$$\begin{aligned} &\forall x \in H(x) \rightarrow M(x) \\ &\forall x \in F(x) \rightarrow H(x) \\ &\forall x \in F(x) \rightarrow M(x) \\ &\exists x_1 \in F(x) \wedge M(x) \end{aligned} \qquad (1)$$

Clearly here the point is to define an ***adjective*** that all elements of the set must present as a necessary property; independently on the subsets they may belong (e.g. a human being is mortal if she/he is a philosopher, a politician or a tennis player).

Let us now look at the following statements:

a)*"X is a red-headed basketball player, all basketball players are persons, and therefore X is red-headed person".*

$$\begin{aligned} &\exists x_1 \in B(x) \wedge R(x) \\ &\forall x \in B(x) \rightarrow H(x) \\ &\exists x_1 \in H(x) \wedge R(x) \end{aligned} \qquad (2)$$

b) *"X is an old basketball player, all basketball players are persons, and therefore X is an old person".*

$$\begin{aligned} &\exists x_1 \in B(x) \wedge O(x) \\ &\forall x \in B(x) \rightarrow H(x) \\ &\exists x_1 \in H(x) \wedge O(x) \end{aligned} \quad (3)$$

c) *"X is a small basketball player, all basketball players are persons, and therefore X is a small person".*

$$\begin{aligned} &\exists x_1 \in B(x) \wedge S(x) \\ &\forall x \in B(x) \rightarrow H(x) \\ &\exists x_1 \in H(x) \wedge S(x) \end{aligned} \quad (4)$$

d) *"X is a good basketball player, all basketball players are persons, and therefore X is a good person".*

$$\begin{aligned} &\exists x_1 \in B(x) \wedge G(x) \\ &\forall x \in B(x) \rightarrow H(x) \\ &\exists x_1 \in H(x) \wedge G(x) \end{aligned} \quad (5)$$

I believe that most readers would agree on the validity of statements a) but very a few would accept statement d) as a correct way of reasoning, although syllogisms from (2) to (5) are all formally correct. In implementing set commensurability, the problem is then to understand which adjectives describing an element can be generalised to the whole set, without any logical or ontological mistake. Geach synthesises this issue very clearly: "Although a tennis stroke or a chess move is a human act, it does not follow that a good tennis stroke or a good chess move is a good human act, … " (Geach, 1956, p. 37).

Here the question is: *when set commensurability is logically possible and correct?* Geach's (1956) distinction between *attributive and predicative adjectives* can help us in answering this question. In Geach's own words: "There are familiar examples of what I call attributive adjectives. Big and small are attributive; *x* is a big flea does not split up into *x* is a flea and *x* is big, nor *x* is a small elephant into *x* is an elephant and *x* is small; for if these analyses were legitimate, a simple argument would show that a big flea is a big animal and a small elephant is a small animal. Again, the sort of adjective that the mediaevals called alienans is attributive; *x* is a forged banknote does not split up into *x* is a banknote and *x* is forged, nor *x* is the putative father of *y* into *x* is the father of *y* and *x* is putative. On the other hand, in the phrase a red book red is a predicative adjective in my sense, although not grammatically so, for is a red book logically splits up into is a book and is red.
I can now state my first thesis about good and evil: good and bad are always attributive, not predicative, adjectives" (Geach, 1956, p. 32).

Even if Geach's arguments were developed in the context of moral philosophy, they have an extraordinary explicative power for evaluation too. In fact, evaluation is all about an action ***a*** being declared better, worse or equal than another action ***b***. However, although Geach saw the clear difference between predicative and attributive adjectives, he only gave examples of them but no general definition was provided. Here I try then to invent one; in my opinion, this way of reasoning can be generalised by defining the new concepts of absolute and relative predicative adjectives.

An adjective is ***absolute predicative*** if its meaning does not change in relation to the subsets considered. It is an intrinsic characteristic of the object considered. The characteristic of being a red-headed person does not change if we consider subsets such as police officers, politicians, scientists or basketball players. In terms of measurement theory (see e.g. Roberts, 1979), an

absolute predicative adjective is always measured on a nominal scale i.e. individual characteristics are grouped into a set of equivalence classes.

Let us know consider the case of an old person, apparently the adjective *old* seems absolute predicative too, but indeed it is not, because one could argue that old actually is dependent on the noun. If economists are, on average older than an average middle aged person, then an old economist is different from an old person. The point becomes clearer if one uses a different profession, e.g. if one refers to "an old basketball player" (statement b) it becomes clear that old in this kind of context can actually have a complete different meaning. The same way of reasoning applies to *small*, which Geach considers an attributive adjective. Clearly, a small elephant is not a small animal and a small basketball player (statement c) is not a small person; the meaning of this adjective changes over different subsets[1], thus in formal logic terms, the law of the excluded middle is lost (since it is not true that e.g. a person is only young or old). An adjective is ***relative predicative*** if it does not hold its meaning once one switches to a larger or different set of objects. It describes a characteristic that is dependent on the relative comparisons among the objects considered. In terms of measurement theory, a relative predicative adjective is always measured on an ordinal scale[2].

An adjective is ***attributive*** if it does not have any meaning when referred to a different set or problem framework. A good person can be a bad basketball player and a good economist can be a bad person. Being good or bad depends also on the notion of quality used, which depends on the use connected to the object to be evaluated. For example, it is hard to defend that a good car is a good mean of transport to travel inside a city's historical centre. I agree with Geach that evaluative adjectives are always attributive. Given a claim that *"x is better than y"* a proper response is *"x is better what than y?"* Similar points can be made about the adjective *"valuable"* and *"is more valuable than".*

At this stage, the following conclusion can be derived: *when considering adjectives,* set commensurability is correct only if the adjectives considered are *absolute predicative* ones. An

---

[1] This type of uncertainty can be modelled by using the concept of fuzzy sets. Fuzzy sets, as formulated by Zadeh (1965), are based on the simple idea of introducing a degree of membership of an element with respect to some sets. Fuzzy uncertainty considers all cases between 0 (non-membership) and 1 (complete membership), and it is represented by means of the membership functions. Let us assume that the symbol *U* means the entire set (Universe of discourse). In classical set theory, given a subset Ω of *U*, each element $x \in U$ satisfies the condition: either *x* belongs to Ω, or *x* does not belong to Ω. The subset Ω is represented by a function $f_{\Omega} : \mathrm{U} \rightarrow [0,1]$. The function f_{Ω} is called a characteristic function of the set Ω. Fuzzy sets are then introduced by generalizing the characteristic function f_{Ω}. Let *U* again be a universe of discourse with $x \in U$. Then a fuzzy set $\tilde{\Omega}$ in *U* is a set of ordered pairs $\left\{x, \mu_{\tilde{\Omega}}(x)\right\}$, $\forall x \in U$ where $\mu_{\tilde{\Omega}} : U \rightarrow \Theta$ is a membership function which maps $x \in U$ into $\mu_{\tilde{\Omega}}(x)$ in a totally ordered set Θ (called the membership set) and $\mu_{\tilde{\Omega}}(x)$ indicates the grade of membership of *x* in $\tilde{\Omega}$. Generally, the membership set is restricted to the closed interval [0, 1]. A fuzzy set is completely determined by its membership function. By using fuzzy sets, a person may belong to the subsets of "old persons", "middle aged" or "young persons" simultaneously, but of course with different degrees of membership. One has to note that although this modelling technique can be useful in practical situations, it does not allow the degree of precision needed by the theoretical concept of set commensurability.

[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 $m: O \rightarrow R$, 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. Let us suppose that there is a measurement procedure or rule for assigning a number *m(o)* to each object $o \in O$, and suppose that the following statements are true for any pair of objects o_1 and $o_2 \in O$:

$$\begin{cases} m(o_1) \neq m(o_2) \; only \; if \;\; t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2) \; only \; if \;\; t(o_1) > t(o_2) \end{cases} \qquad (6)$$

Any measurement procedure for which equation (6) applies is an example of *ordinal scaling*, or measurement at the ordinal level. Thus in comparing two basketball players *x* and *y*, we can state that *x* is smaller or older than *y*, but we cannot state anything about a third person *z*, with whom no pair-wise comparison was made.

adjective *Z* is absolute predicative if it passes the *ontological* check of the two following logical tests:

(1) $if\ \exists x_1 \in Z(x) \wedge W(x) \Leftrightarrow x_1 \in Z(x) \wedge x_1 W(x)$

(2) $if\ \exists x_1 \in Z(x) \wedge W(x)\ and\ if\ \forall x \in W(x) \rightarrow V(x)\ then\ \exists x_1 \in Z(x) \wedge V(x)$

In plain language, test (1) implies statements such as *"if x_1 is red and it is a car then x_1 is a red car*" and test (2) *"if x_1 is a red car and all cars are a mean of transport then x_1 is a red mean of transport".* Adjectives that fail such tests are relative predicative or attributive adjectives which always imply weak comparability based on incommensurability. For example, the adjective "good" clearly fails (2), statements such as *"x_1 is a good car, all cars are a mean of transport, and therefore x_1 is a good mean of transport"* or *"x_1 is a good economist, all economists are human beings, and therefore x_1 is a good human being"* are invalid arguments on the light of a real-world corroboration.

If evaluative adjectives like "good" and "valuable" are attributive in standard uses, that does not however preclude the possibility of rational choices between objects which do not fall into the range of a single comparative. ***Weak comparability*** based on ***incommensurability*** is compatible with the existence of such limited ranges; for example, urban quality of life is not evaluated as good or bad as such, but rather, as good, bad, beautiful or ugly in relation to different descriptions or indicators. 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 use of these value terms in such contexts is attributive clearly.

Now let us go back to the case of relative predicative adjectives. One could argue that isn't it true that there can be commensurability even where relative predicative adjectives are involved as long as one sticks to a single measurement unit? So, going back to the example of big and small as relative predicative adjectives, the relative sizes of elephants and fleas are indeed commensurable as long as one considers a single measurement unit such as kilos, (or pounds), centimetres (or inches) in diameter and so on. The point here is then that *adjectives* are relative predicative but the ***property*** behind this type of adjective can in principle be measured exactly on a quantitative measurement scale[3]. ***Factual rod commensurability*** is based on this attempt to look for *one existing property* common to all objects to be compared and measurable by using one

[3] 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 (7) is true:

$$\begin{cases} m(o_1) \neq m(o_2)\ only\ if\ \ t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2)\ only\ if\ \ t(o_1) > t(o_2) \\ t(o) = x\ iff\ m(o) = ax + b, where\ a \in R^+ \end{cases} \quad (7)$$

That is, the numerical measurement *m(o)* is some *affine function* of the true magnitude *x*. When (7) 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 (8) true:

$$\begin{cases} m(o_1) \neq m(o_2)\ only\ if\ \ t(o_1) \neq t(o_2) \\ m(o_1) > m(o_2)\ only\ if\ \ t(o_1) > t(o_2) \\ t(o) = x\ iff\ m(o) = ax, where\ a \in R^+ \end{cases} \quad (8)$$

When the measurement operation defines a function such as the statement contained in (8), 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.

measurement unit; obviously comparison of objects is possible according to the characteristics of this property only. In a syllogistic form, the apples and oranges example could be put as follows: *All consumer's goods have a price on the market, all fruits on the market are consumer's goods, then all fruits on the market have a price, apples and oranges are fruits on the market, thus they have a price by which can be compared*. By using formal logic it is:

$$
\begin{aligned}
&\forall x \in C_M(x) \rightarrow P(x) \\
&\forall x \in F_M(x) \rightarrow C_M(x)) \\
&\forall x \in F_M(x) \rightarrow P(x) \\
&\forall x \in A(x) \rightarrow F_M(x) \\
&\forall x \in O(x)) \rightarrow F_M(x) \\
&\exists x_1 \in A(x) \wedge P_1(x) \\
&\exists x_2 \in O(x) \wedge P_2(x) \\
&P_1(x) > P_2(x) \cup P_1(x) = P_2(x) \cup P_1(x) < P_2(x)
\end{aligned}
\qquad (9)
$$

There is no doubt that rod commensurability is very attractive, in its framework it is possible to make statements ontologically and logically correct on a quantitative scale of measurement and a complete pre-order among objects to be evaluated can always be derived.

In economic terms a question immediately arises: is the search for the *"common rod of money"* useful? I do think it is very useful, but one has to have clear that money values are worth to be used when they are connected to *one value* and *one institution* only, i.e. economic efficiency and markets. *They fail to incorporate other objectives and values.*

Monetary valuation methods are based on phenomena such as consumer's surpluses, market failures, demand curves which are just a partial point of view, since connected with one institution and one value only: *markets and efficiency.* 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. I believe that the point is not to be against giving economic value to natural resources, to human health (or even lives) or to cultural heritage. When one wishes to preserve 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[4] 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*. 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 the travel cost method or contingent valuation). These are different types of value. *The point is that social decisions involve multiple types of values, of which economic efficiency is only one. Therefore it is misleading to make social decisions based only on that one value*.

The classical Adam Smith's example on the value of diamonds versus water is relevant here. No doubt in a city environment everyone would prefer diamond over water, however in a different environment, e.g. a boat in the middle of the ocean, water has definitely a higher value than diamonds. Economic values depend on subjective human preferences, no discussion about this. Attempts to explain economic values through objective, context invariant categories such as

[4] Although money values can be considered a form of *factual rod commensurability*, it is important to remember that to compute total economic values has nothing to do with the idea of a *"true"* or *"correct"* value. All monetary valuation attempts suffer deep philosophical problems (see e.g. Hansson, 2007; Sagoff, 1988) and technical uncertainties (see e.g. Grüne-Yanoff, 2009; Hansen, 2011; Martinez-Alier *et al.,* 1998; 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?

energy are an obvious non-sense. On the other side, e.g. Odum's Emergy[5] measures (Odum, 1996) can be a good proxy of the ecological value of an ecosystem. Galapagos Islands have a higher ecological value than the Dutch Inside Sea surely, but the same does not necessarily apply to economic value (economic indeed would favour the Inside Sea, which, since totally eutrophised, offers an important economic service receiving all the nutrients coming from human activity). Different values, since they are related to different objectives and institutions, cannot be merged into only one metric, this is a very simple truth.

It is interesting to note that the consistency between measurement units and objectives applies even when two metrics are equivalent. A famous example is the one of energy and labour valuation in the Marxist theory of value (Martinez-Alier, 1987). Podolinski, an Ukrainian medical doctor, arrived at the conclusion that a human being in average can devote to work about 500 Kcal a day. This means that, by considering a human being as a thermal machine, 1/5 of Kcal coming from the food can be transformed into muscular effort, i.e. mechanical energy. Both Marx and Engels agreed with Podolinski's argument that labour theory of value would be equivalent to energy theory of value, but they strongly rejected the possibility of substituting labour valuation with energy valuation. The reason is obvious, their *ideological objective* was to show that in a production process, there is exploitation of the workers by the capital; this message was easy to communicate in the framework of labour theory but totally unintelligible if an energy theory of value would be used, since the economy/society relationship is excluded completely.

Potential rod commensurability has two distinguishing characteristics:
1) It is based on the search for one plausible common property among the objects to be evaluated, although this property *is not necessarily recognizable in the real world.*
2) This common property tries *to represent many dimensions simultaneously*.

An example is the so-called ecological footprint sustainability index (Wackernagel and Rees, 1995). The peculiar characteristic of this index is that it starts from the assumption that every category of energy and material consumption and waste discharge requires the productive or absorptive capacity of a finite area of land or water. If one sums up the land requirements for all categories of consumption and waste discharge of a defined population, the total area represents the ecological footprint of that population. The obvious problem here is that while in comparing apples and oranges, one can know their content in Kcal or their price, how can one know how much space is needed for their production? This is not an observable property; in this framework, the solution of the representation problem of measurement theory implies the use of scientifically sound conversion factors that can transform *all various dimensions into land.* These conversion factors necessarily depend on arbitrary assumptions on e.g. agricultural production system, biological productivity, characteristics of soil, age of trees, and so on. From an evaluation point of view, it is interesting to note that an intensive agricultural production system in this case is better than an extensive one. In fact energy based systems reduce the virtual space occupied by e.g. a city, but unfortunately, at the same time imply much more CO_2 emissions and loss of biodiversity due to the use of fertilizers and pesticides and the introduction of exotic species.

One should observe that from a policy point of view, this kind of commensurability is very risky; since a multidimensional concept is compressed into a single dimension only, no consistency between policy objectives and the metric used exists anymore. For example, if sustainability is summarised into the ecological footprint, a rational policy would be to transform the Coliseum into a wooded area since this option decreases the ecological footprint of Rome! Clearly this index is an example of ecological reductionism where e.g. socio-economic and cultural aspects are completely neglected. Indeed in my opinion, the objective of this index is not policy, but communication. In fact its relative popularity is due to the fact that as a pedagogical tool, it is very understandable (space in terms of footprint) and its message very clear: humans should reduce their impact on the planet by changing their life style. Like in the case of Marx theory of value, the objective linked to the metric chosen is ideological.

[5] *Emergy* is the "available solar energy used up directly and indirectly to make a service or product" (Odum, 1996, p. 8).

Another illustration of potential rod commensurability is the use of shadow prices in macroeconomic environmental accounts. Here the objective is pragmatic, i.e. to account for environmental goods and services that have no price on the market. Prices are used not to implement efficiency but simply to aggregate different items of the so-called environmental capital (see e.g. Dasgupta, 2001; Pearce and Turner, 1990; World Bank, 2011). According to this scholar tradition, measurements of natural capital stock made exclusively in physical terms are problematic because of the difficulty in adding up different physical quantities expressed in different units. On the contrary, by valuing each resource stock in money terms, the total value of natural capital can be measured[6]. One obvious problem here is that many natural resources (e.g., air, water and wilderness) do not have observable prices. Even those prices that do exist may not be useful; they may be affected by market imperfections and taxes, and they may exclude externalities involved with the production and use of the resource. Thus one would need to find implicit or shadow prices in some way.

Shadow prices have been invented for implementing the concept of "real opportunity cost" when a market for a good or service does not exist (e.g. the provision of many public goods), thus the ratios of shadow prices are marginal social rates of substitution among the various capital assets[7]. From a pure technical point of view, it is important to remember that shadow prices are primarily meant to implement efficiency, i.e. resource prices reflect conditions at the margin, and thus this is their natural objective. In the framework of sustainability, this may give rise to counterintuitive results. Let us imagine that the monetary value of a country natural capital stock has been properly measured and its value is known to policy-makers. When considering flows from a natural resource stock, where the resource flow is the product of price and quantity used, when quantity declines the corresponding shadow price unavoidably rises over time (being a function of its relative scarcity), at the same rate or faster than the rate of decrease in the physical stock of the natural resource. This implies that the value over time of the natural capital stock remains constant or may even rise, while the physical stock is declining; but unfortunately this physical scarcity will not be recognised by policy-makers since they have information on the value of the natural capital stock.

In summary, the point is that different metrics are linked to different objectives and values. To use only one measurement unit for incorporating a plurality of dimensions, objectives and values, implies reductionism necessarily.

3. Implications of incommensurability for public policy: from welfare economics to well-being economics

The world is characterised by deep *complexity*. This obvious observation has important implications on the manner in which policy problems are represented and decision-making is framed. Various authors claim that modern public economic policy needs to expand its empirical relevance by introducing more and more realistic (and of course more complex) assumptions in its models. In this context, one of the most interesting research directions in contemporary public economics, is the attempt of taking into account political constraints, interest groups and collusion effects explicitly (see e.g. Laffont, 2000, 2002; van Winden, 1999), as a consequence, *transparency* becomes an essential feature of public policies (Stiglitz, 2002). This implies that to reach a ranking of policy options, there is a previous need for deciding about *what is important* for

[6] In terms of history of the economic thought, it is worthy to remember that the issues concerning the nature, role and measurement of capital goods was the focus of the so-called Cambridge controversy (Harcourt, 1972). Although the term capital was referred to artificial capital, the results can be extended to natural capital too, if money valuation is used, in particular the problem of circularity among the quantity of capital, its monetary value and the rate of interest.

[7] Indeed shadow prices were invented as a solution to the debate on economic calculus in a socialist economy. Hayek, replying to Neurath wrote (1935, p. 31): "Neurath was quite oblivious of the insuperable difficulties which the absence of value calculations would put in the way of any rational economic use of the resources...". Or, as Von Mises had put it (Von Mises, 1920, in Hayek, ed. 1935, p. 111), "Where there is no free market, there is no pricing mechanism; without a pricing mechanism, there is no economic calculation". Kantorovich (1939) found a possible solution to this dilemma by inventing shadow prices.

different social actors as well as *what is relevant* for the representation of the real-world entity described in the model.

In the framework of public policy, it is important to further distinguish the concepts of social incommensurability and technical incommensurability (Munda, 2004). ***Social incommensurability*** refers to the existence of a multiplicity of legitimate values in society. Any social decision problem is characterised by conflicts between competing values and interests and different groups and communities that represent them. Choosing any particular operational definition for *value* and its corresponding valuation technique involves making a decision about what is important and real. Distributional issues play a central role (Olson, 1982). Any policy option always implies winners and losers, thus it is important to check if a policy option seems preferable just because some dimensions (e.g. the environmental) or some social groups (e.g. the lower income groups) are not taken into account.

Accepting low values for a negative externality that provokes an impact on poor community is a "political decision", far from being ethically neutral. Some years ago, an internal document of the World Bank, subsequently made public, suggested that toxic waste should be located in Africa, since the cost of the compensation was extremely low and therefore such solution has to be considered as the most efficient one. One should note that the issue of *value free* Science is a key issue for real-world policy and not a philosophical debate only. For example, an influential economist claimed that his work for the intergovernmental Panel on Climate Change (IPCC), where lives of people in rich countries are valued up to fifteen times higher than those in poor countries, was a matter of *scientific correctness* versus *political correctness.* (New Scientist, 19 August, 1995). Is it really a matter of value free scientific correctness to use valuations based on assessments of a community's willingness and ability to pay to avoid risks of death[8]? I believe that both Kaldor and Hicks would answer no to this question. They stated clearly: "… it is quite impossible to decide on economic grounds what particular pattern of income-distribution maximises social welfare" (Kaldor, 1939, p. 551); "… there will be an indefinite number of different possible optima, distinguished from one another by differences in the distribution of social wealth" (Hicks, 1939, p. 701).

The new nature of the problems faced in this third millennium (e.g., food security, genetic modified organisms, climate change, …), implies that very often, when deciding on problems that may have long term consequences, we are confronting issues "*where facts are uncertain, values in dispute, stakes high and decisions urgent*" (Funtowicz and Ravetz, 1991). An obvious question is then who has the power to simply complexity? In this case, scientists cannot provide any useful input without interacting with the rest of society and the rest of the society cannot perform any sound decision making without interacting with the scientists. That is, the question of "how to improve the quality of a policy process" must be put, quite quickly, on the agenda of "scientists", "policy makers" and indeed the whole society. This extension of the "peer community" is essential for maintaining the quality of the process of policy-making when dealing with real-world complex systems. The fact that "one's welfare economics will inevitably be different according as one is a liberal or a socialist, a nationalist or an internationalist, a christian or a pagan" (Hicks, 1939, p. 696) is the normal state of affairs in policy decisions. *I do not see any 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 from one's preferences as a citizen significantly. In my opinion, it is much more scientific an approach which deals with such a plurality of values than one which solve all conflicts by imposing a perspective considered superior on some ethical or technical grounds.

[8] 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 political one instead.

Technical incommensurability comes from the multidimensional nature of policy issues. The existence of different dimensions, levels and scales at which a hierarchical system can be analyzed implies the unavoidable existence of non-equivalent descriptions of it. As discussed by Giampietro (2003) even a simple "objective" description of a geographical orientation is impossible without taking an arbitrary subjective decision on the system scale considered relevant. In fact the same geographical place, e.g., in the USA, may be considered to be in the north, south, east or west according to the scale chosen as a reference point (the whole USA, a single state and so on)[9]. One should note that the construction of a descriptive model of a real-world system depends on very strong assumptions about (1) the *purpose* of this construction, e.g. to evaluate well-being or sustainability (2) the *scale* of analysis, e.g. a city, a region or a country and (3) the set of *dimensions, objectives* and *indicators* used for the evaluation process. A reductionist approach for building a descriptive model can be defined as the use of just one measurable indicator (e.g. GDP per capita), one dimension (e.g. economic), one scale of analysis (e.g. region), one objective (e.g. the maximisation of economic efficiency) and one time horizon. If one wants to avoid reductionism, there is a clear need to take into account incommensurable dimensions using different scientific languages coming from different legitimate representations of the same system. This is what Neurath (1973) called the need for an *"orchestration of sciences"* (advocacy of interdisciplinarity and plurality of visions and judgement can also be found in Colander, 1994 and Hansen, 2011).

In summary, I think that instead of focusing on "missing markets" as a source of theoretical and empirical problems, or trying to explain economic values by means of energy or other common rod measures (clearly a non-sense from an economic point of view), we should focus on the creative power that missing markets have, because they push us away from commensurability towards a multidimensional evaluation of evolving realities[10] implementing the incommensurability principle. This is the reason why I prefer to use the term *Well-Being Economics* instead of "*Welfare Economics of Well-Being"* suggested by Arrow *et al.,* 2012.

I believe we can accept as true the statement that incommensurability does not imply incomparability; on the contrary incommensurability is the only rational way to compare various objects under different methodological assumptions than maximisation or optimisation (Sen, 1997, 2000; Sen and Williams, 1982). It is in terms of weak comparability that evaluation has to take place in practice. Evaluation of objects relative to different descriptions invokes not just different practices and perspectives, but also the different criteria and standards for evaluation associated with these. It presupposes value-pluralism. This is exactly the basic idea of multi-criteria evaluation, which can be considered a form of applied consequentialism[11]. Weak comparability can therefore be implemented by using multi-criteria evaluation. In empirical evaluations of public projects and public provided goods, multi-criteria evaluation seems to be an adequate policy tool since it allows taking into account a wide variety of evaluation criteria (e.g. environmental impact, distributional equity, and so on) and not simply profit maximisation, as a private economic agent would mainly do.

The basic idea of multi-criteria evaluation is that in evaluation problems, we have first to establish *objectives*, i.e. the direction of the desired changes of the world (e.g. maximise profits, minimise environmental impact, minimise social exclusion, etc.) and then find useful practical indicators (called criteria). *The evaluation criterion* is the basis for evaluation in relation to a given objective. It is a *function* that associates each alternative with a variable indicating its desirability according to expected consequences related to the same objective, e.g. GDP, saving rate and inflation rate inside the objective "growth maximization". *The criterion score* is a constructed measure stemming from a process that represents, at a given point in space and time, a shared

[9] These multiple-identity/multiple-scale systems can be defined as "*Learning Holarchies"*. A *"holon"* is a whole made of smaller parts (e.g. a human being made of organs, tissues, cells, atoms) and at the same time it forms a part of a larger whole (an individual human being is a part of a household, a community, a country, the global economy) (see Koestler, 1969).

[10] "There is great pressure for research into techniques to make larger ranges of social value commensurable. Some of the effort should rather be devoted to learning - or learning again, perhaps - how to think intelligently about conflicts of value which are incommensurable" (Williams, 1972, p. 103). A call for dealing explicitly with incommensurability can also be found in Arrow (1997).

[11] Here I disagree with Hansson (2007, p. 163) who considers cost-benefit analysis" the only well-developed form of applied consequentialism".

perception of a real-world state of affairs consistent with a given criterion. To give an example, when comparing two countries, within the economic dimension, one objective could be "maximization of economic growth"; the criterion might be R&D performance, the criterion score could be "number of patents per million of inhabitants". ***Valuation*** can then be referred to the process by which a criterion score is constructed. Since in general, objectives are in conflict, multi-criteria mathematical aggregation rules look for so-called *compromise solutions*.

In formal terms discrete multi-criteria evaluation problems can be described in the following way (Arrow and Raynaud, 1986; Roy, 1996; Figueira et al., 2005): *A* is a finite set of *N* feasible options; *M* is the number of different points of view or evaluation criteria g_m $m=1, 2, \ldots, M$ considered relevant in an evaluation problem, where the option ***a*** is evaluated to be better than option ***b*** (both belonging to the set *A*) according to the *m-th* point of view if $g_M(a)>g_M(b)$. This information can be synthesised in a matrix called *evaluation or impact matrix* (see Table 1).

		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 summary, the information contained in the impact matrix useful for solving[12] 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.

In 1986 Kenneth Arrow and Hervé Raynaud published a very influential book titled *"Social choice and multicriterion decision-making",* where the formal analogies between the discrete multi-criterion problem and the social choice one are analysed deeply. This book is based on the assumption that, 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 decision theory fully (at least when no intensity of preference and no indifference/preference thresholds[13] are used; for a

[12] 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.

[13] By introducing a positive constant indifference threshold *q* the resulting preference model is the *threshold model:*

$$\left\{\begin{array}{ll} a_j P a_k & \Leftrightarrow \; g_m(a_j) > g_m(a_k)+q \\ a_j I a_k & \Leftrightarrow \; \left|g_m(a_j) - g_m(a_k)\right| \le q \end{array}\right\}$$

where a_j and a_k belong to the set *A* of alternatives and g_m to the set *G* of evaluation criteria.
The famous bold paradox in Greek philosophy (how many hairs one has to cut off to transform a person with hairs to a bold 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 elements, or refuses to

recent discussion of these technical issues see Munda, 2012a,b). However in my opinion, the relations between social choice and multi-criteria evaluation are stronger than the simple mathematical analogy. In fact I consider that multi-criteria evaluation is a type of applied democracy when it is used for evaluating policy options, this is the main idea behind *Social Multi-Criteria Evaluation (SMCE)* (Munda, 2004, 2008).

As a tool for conflict management, SMCE has demonstrated its usefulness in many policy problems in various geographical and cultural contexts (see e.g. Gamboa, 2006; Garmendia and Stagl, 2010; Monterroso *et al.,* 2011; Munda and Russi, 2008; Özkaynak, 2008; Scolobig *et al.,* 2008; Soma and Vatn, 2009; Straton *et al*., 2010; Zendehdel *et al.,* 2010). The main point of force 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. From this point of view, social multi-criteria evaluation can be considered as a tool for implementing political democracy. Social multi-criteria evaluation puts its emphasis on the transparency issue; the main idea being that results of an evaluation exercise depends on the way a given policy problem is structured and thus the assumptions used, the ethical positions taken, and the interests and values considered have to be made clear. In this framework, mathematical models still play a very important role: the one of guaranteeing consistency between assumptions used and results obtained.

Finally, one should note that these considerations apply at a macro level of analysis too. Since well-being is a multidimensional concept, in evaluating the performance of a region, country or even a continent, a plurality of indicators is needed. Often, some indicators improve while others deteriorate (Nardo *et al.,* 2008). This is the classical conflictual situation dealt with in multi-criteria evaluation. In a macroeconomic framework too, it is essential to identify how countries or regions improve or decline under certain assumptions, and to help the framing of the debate around the conceptual framework used, i.e. which representation of reality (and thus which societal values and interests) has been taken into consideration (for a technical treatise see e.g. Munda, 2005; Munda and Nardo, 2009).

4. Conclusions

In proportion to a more or less complete acceptance of the arguments presented here, the following conclusions can be drawn:

1. Well-being is a multidimensional concept, thus a key research topic is on how we can measure and evaluate this plurality of dimensions for policy decisions. In the literature two main approaches exist: strong and weak comparability.

2. Commensurability, a necessary condition for strong comparability, can be implemented by means of “set commensurability” and “rod commensurability”. Both of them are not of a general applicability. Set commensurability can be applied only when absolute predicative adjectives are considered. Potential rod commensurability is very risky since it attempts to “invent” a plausible common property which synthesises many dimensions at the same time. Factual rod commensurability is always implied by a valuation exercise; it is desirable but one has to remember that e.g. monetary valuation methods are based on 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. It is misleading to take social decisions based on only one type of value.

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 characterisations of preference modelling with thresholds can be found in Roubens and Vincke (1985).

3. Monetary valuation techniques are the only ones that can answer these two questions: 1) How many resources society is willing to devote to a given objective? 2) How much society has to pay for compensation after e.g. an accident? Their desirability in this context is not questioned here.

4. Different metrics are linked to different objectives and values. To use only one measurement unit for incorporating a plurality of objectives and values, implies reductionism necessarily.

5. Weak comparability of values, which is grounded on incommensurability, can be considered the main methodological foundation of evaluation and policy in the framework of well-being economics. Incommensurability does not imply incomparability; on the contrary incommensurability is the only rational way to compare societal options under a plurality of policy objectives.

6. Weak comparability can be implemented by using multi-criteria evaluation, which is a formal framework for applied consequentialism under incommensurability. Social Multi-Criteria Evaluation, in particular, allows considering both technical and social incommensurability simultaneously.

7. We can define ***evaluation*** as the combination of *representation* (social actors, criteria, weights and actions considered), *valuation* (construction of criterion scores), *mathematical aggregation* (properties of the algorithms used) and *quality check* (transparency of the steps by which both technical and social commensurability have been tackled) connected to a given policy problem.

ACKNOWLEDGEMENTS: Comments by Michela Nardo on previous drafts of this paper are gratefully acknowledged. This research has been partially developed in the framework of the Spanish Government financially supported project SALMON (HAR2010-20689-02-01).

References

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

Arrow (1997) - Invaluable Goods, *Journal of Economic Literature*, Vol. 35, No. 2, pp. 757- 763.

Arrow K. J., Dasgupta P., Goulder L. H., Mumford K. J. and Oleson K. (2012) - Sustainability and the measurement of wealth, *Environment and Development Economics* / Volume 17 / Issue 03 / pp. 317 – 353.

Chang R. (ed.) (1997) - *Incommensurability, Incomparability, and Practical Reason*, Cambridge: Harvard University Press.

Colander D. (1994) - Vision, judgment, and disagreement among economists, *Journal of Economic Methodology*, Volume 1, Issue 1, pages 43-56.

Dasgupta P. (2001) – Valuing objects and evaluating policies in imperfect economies, *Economic Journal,* 111 (May), pp. C1–C29.

Duclos J.Y., Sahn D.E. and Younger S.D. (2006) – Robust multidimensional poverty comparisons, *Economic Journal,* 116 (October), pp. 943–968.

Easterlin, R. A. (1995) - `Will raising the incomes of all increase the happiness of all?' *Journal of Economic Behavior and Organization*, vol. 27, pp. 35-47.

Easterlin R.A. (2001) – Income and happiness: towards a unified theory, *Economic Journal,* 111 (July), pp. 465-484.

Figueira, J., Greco, S. and Ehrgott, M. (eds.) (2005) *Multiple-criteria decision analysis. State of the art surveys*. Springer International Series in Operations Research and Management Science, New York.

Frey B.S. and Stutzer A. (2002) - What can economists learn from happiness research?, *Journal of Economic Literature*, Vol. XV, (June), pp. 402-435.

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.

Gamboa G. (2006) - Social multi-criteria evaluation of different development scenarios of the Aysén region, Chile, *Ecological Economics*, Volume: 59, Issue: 1, Pages: 157-170

Garmendia, E. , **Stagl, S.** (2010) - Public participation for sustainability and social learning: Concepts and lessons from three case studies in Europe, *Ecological Economics,* Volume 69, Issue 8, 15, Pages 1712-1722
Geach P. T. (1956) – Good and Evil, *Analysis*, Vol. 17, pp. 32–42 - Reprinted in *Foot, Philippa* (ed.) (1967). *Theories of Ethics*, Oxford University Press, Oxford, pp. 64–73.
Giampietro, M. (2003) - *Multi-Scale Integrated Analysis of Agroecosystems*, CRC Press, New York.
Grüne-Yanoff T. (2009) - Mismeasuring the value of statistical life, *Journal of Economic Methodology*, Volume 16, Issue 2, pages 109-123.
Hayek, F. A. (ed.) (1935) - *Collectivist Economic Planning.* Reprinted Augustus M. Kelley, New York, 1970.
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.
Harcourt G. C. (1972) - *Some Cambridge Controversies in the Theory of Capital*, Cambridge University Press, Cambridge.
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.
Kantorovich L.V. (1939) - Mathematical Methods of Organizing and Planning Production, Leningrad University Press, published in English in *Management Science,* Vol. 6, No. 4 (July, 1960), pp. 366–422.
Koestler A. (1969) – Beyond atomism and holism: the concept of the holon, in A. Koestler and J.R. Smythies (eds.) – *Beyond reductionism*, Hutchinson, London, pp. 192-232.
Laffont J.J. (2000) – *Incentives and political economy*, Oxford University Press, Oxford.
Laffont J.J. (2002) – Public economics yesterday, today and tomorrow, *Journal of Public Economics*, 86, pp. 327-334.
Luce R.D. (1956) – Semiorders and a theory of utility discrimination, *Econometrica,* 24, pp.178-191.
Martinez-Alier, J. (with Klaus Schluepmann) (1987) - *Ecological Economics.* Blackwell, Oxford.
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.
Michalos A.C. (1980) - Satisfaction and happiness, *Social Indicators Research,* Volume 8, Issue 4, pp 385-422.
Michalos A.C. (1997) - Combining social, economic and environmental indicators to measure sustainable human well-being, *Social Indicators Research,* Volume 40, Issue 1-2, pp 221-258.
Monterroso, I. , Binimelis, R., Rodríguez-Labajos, B. (2011) - New methods for the analysis of invasion processes: Multi-criteria evaluation of the invasion of Hydrilla verticillata in Guatemala, *Journal of Environmental Management,* Volume 92, Issue 3, Pages 494-507
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. (2005) - "Measuring sustainability": a multi-criterion framework. *Environment, Development and Sustainability* Vol 7, No. 1, pp. 117-134.
Munda G. (2008) – *Social multi-criteria evaluation for a sustainable economy*, Springer, Heidelberg, New York.
Munda G., Russi D. (2008) - Social multi-criteria evaluation of conflict over rural electrification and solar energy in Spain, *Environment and Planning C: Government and Policy* Volume 26, pp. 712-727.
Munda G. and Nardo M. (2009) - Non-compensatory/non-linear composite indicators for ranking countries: a defensible setting*, Applied Economics* Vol. 41, pp. 1513-1523.
Munda G. (2012a) – Choosing aggregation rules for composite indicators, *Social Indicators Research*, Volume 109, Issue 3, pp. 337-354,
Munda G. (2012b) – Intensity of Preference and related Uncertainty in Non-Compensatory Aggregation Rules, *Theory and Decision,* Volume 73, Issue 4, pp. 649-669.
Nardo, M., Saisana, M., Saltelli, A., Tarantola, S., Hoffman, A., Giovannini, E. (2008) -*Handbook on Constructing Composite Indicators: Methodology and User Guide.,* Paris: OECD Statistics Working Paper.
Neurath O. (1973) - *Empiricism and sociology*. Reidel, Dordrecht.
Odum, H. T. (1996) - Environmental accounting: EMERGY and environmental decision making, Wiley, New York.
Olson, M. (1982) - *The rise and decline of Nations: Economic growth, stagflation and social rigidities*, Yale University press, New Haven, Connecticut.
O'Neill, J. (1993) - *Ecology, Policy and Politics*. Routledge, London.
Özkaynak, B. (2008) – Globalisation and local resistance: Alternative city developmental scenarios on capital's global frontier-the case of Yalova, Turkey, *Progress in Planning,* Volume 70, Issue 2, Pages 45-97.
Pearce D.W., Turner K.R. (1990) - *Economics of natural resources and the environment,* Harvester Wheatsheaf, New York.
Poincaré H. (1935) – *La valeur de la science*, Flammarion.
Raz, J. (1986) - *The Morality of Freedom*, Oxford: Clarendon Press.
Rabinowicz W. (2012) - Value Relations Revisited, *Economics and Philosophy*, 28, pp. 133-164.
Roberts, F. S. (1979) - *Measurement theory with applications to decision making, utility and the social sciences*, Addison-Wesley, London.
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.
Scolobig, A. , Broto, V.C. , Zabala, A. (2008) - Integrating multiple perspectives in social multicriteria evaluation of flood-mitigation alternatives: The case of Malborghetto-Valbruna, *Environment and Planning C: Government and Policy* , Volume 26, Issue 6, Pages 1143-1161.

Sen, A. (1979) - Personal utilities and public judgment: or what's wrong with welfare economics?, *Economic Journal,* vol. 89 (355), pp. 537–58.
Sen, A. (1985) - *Commodities and Capabilities*, Amsterdam: North-Holland.
Sen A. (1997) - Maximization and the act of choice, *Econometrica*, 65, pp. 745-779.
Sen A. (2000) - Consequential evaluation and practical reason, *Journal of Philosophy*, 98, pp. 477-502.
Sen, A. and B. Williams, (eds.) (1982) - *Utilitarianism and Beyond*, Cambridge: Cambridge University Press.
Soma, K. , Vatn, A. (2009) - Local democracy implications for coastal zone management-A case study in southern Norway, *Land Use Policy,* Volume 26, Issue 3, Pages 755-762
Stiglitz J. E. (2002) – New Perspectives on public finance: recent achievements and future challenges, *Journal of Public Economics,* 86, pp. 341-360.
Stiglitz J., Sen A. and Fitoussi J.P. (2009) – *Report by the Commission on the measurement of economic performance and social progress,* available at http://www.stiglitz-sen-fitoussi.fr/en/index.htm
Straton, A.T. , Jackson, S. , Marinoni, O. , Proctor, W., Woodward, E. (2010) - Exploring and Evaluating Scenarios for a River Catchment in Northern Australia Using Scenario Development, Multi-criteria Analysis and a Deliberative Process as a Tool for Water Planning, *Water Resources Management*, Volume 25, Issue 1, Pages 141-164
van Winden F. (1999) – On the economic theory of interest groups: towards a group frame of reference in political economics, *Public Choice*, 100, pp. 1-29.
Vatn A., **Bromley D.W.** (1994) – Choices without prices without apologies, *Journal of Environmental Economics and Management,* 26, pp. 129-148.
Wackernagel, M., Rees, W. E. (1995) - *Our ecological footprint: Reducing human impact on the earth.* Gabriola Island, BC and Philadelphia, PA: New Society Publishers.
Williams B. (1972) - *Morality*, Cambridge University Press, Cambridge.
World Bank (2011) - *The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium*, Washington, DC: World Bank.
Zadeh L.A. (1965) - Fuzzy sets, *Information and Control,* 8, pp. 338-353.
Zendehdel, K. , Rademaker, M., De Baets, B., Van Huylenbroeck, (2010) - Environmental decision making with conflicting social groups: A case study of the Lar rangeland in Iran, *Journal of Arid Environments,* Volume 74, Issue 3, Pages 394-402.