EASY-ECO Vienna Conference 2008
11-14 March
“Governance by Evaluation: Institutional Capacities and Learning for Sustainable Development”

Incorporating stakeholders’ preferences for ex ante evaluation of energy and climate policy interactions: Development of a Multi Criteria Analysis weighting methodology

Grafakos S[a1]., Zevgolis D[b]., and Oikonomou V[c]

[a] *Laboratory of Industrial and Energy Economics (LIEE), School of Chemical Engineering, National Technical University of Athens(NTUA), Greece*
and
Energy Research Centre of The Netherlands (ECN), Unit of Policy Studies, Amsterdam, The Netherlands
[b] *EPU-NTUA, Decision Support Systems Laboratory of ICCS-NTUA, National Technical University of Athens, Greece*
[c] *SOM, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands*

Abstract

Evaluation of energy and climate policy interactions is a high complex issue which has not been addressed systematically. Multi Criteria Analysis (MCA) evaluation processes have been applied widely to different policy and decision cases as they have the ability to cope with high complexity, by structuring and analyzing the policy problem in a transparent and systematic way. Criteria weights elicitation techniques are developed within the framework of MCA to integrate stakeholders’ preferential information in the decision making and evaluation process. There are numerous methods to determine criteria weights which can be used in various ways for different policy evaluation purposes. During decision making, relevant stakeholders and policy makers implicitly or explicitly express their relative importance between the evaluation criteria by assigning weighting factors to them. More particular, climate policy problems lack a simple, transparent and structured way to incorporate stakeholders’ views and values. In order to incorporate stakeholders’ weighting preferences into an ex ante evaluation of climate and energy policy instruments interaction, an integrative constructive weighting methodology has been developed.

This paper presents the main characteristics of evaluation of energy and climate policy interactions, the reasoning behind the development of the weighting tool, its main theoretical and functional characteristics and the results of its testing application to obtain and incorporate stakeholders’ preferences on energy and climate policy evaluation criteria.

[1] Contact person: email address grafakos@central.ntua.gr, grafakos@ecn.nl

The weighting method that has been elaborated to derive stakeholders' preferences for criteria weights is a combination of pair wise comparisons and ratio importance weighting methods. Initially introduces the stakeholders to the evaluation process through a warming up holistic approach for ranking the criteria and then requires them to express their ratio relative importance in pair wise comparisons of criteria by providing them an interactive mean with verbal, arithmetic and visual representation of their preferences. An excel aided tool has been developed for this purpose to facilitate the user (stakeholder) to intervene to the decision making and evaluation process and set his preferences that determine the outcome of policy interactions evaluation. The developed tool has been tested at an experiment examining how stakeholders' preferences and objectives can be derived towards different evaluation climate change and energy evaluation criteria. The paper concludes by stating users' evaluation responses on the application of the methodology.

Key words: energy and climate policy interaction, evaluation, stakeholders' preferences, weighting, multi - criteria analysis

1. Introduction

Energy forms a central element in discussions on sustainable development. First of all, the use of energy supports economic development. Securing affordable energy supply is an important element in the energy policies of many countries. At the same time fossil fuels combustion is the single most important cause of greenhouse gas emissions. Climate change appears to be among the most prominent sustainability problems of this century. Policy decisions and evaluations for climate change and energy often face a high degree of complexity and multi dimensionality that characterizes these types of policy problems.

Kyoto Protocol developed a broad range of policy instruments in order to help Annex I countries to achieve the mitigation of GHG emissions commitments. These instruments often interact between each other and they influence their effectiveness. They can be overlapping, indifferent or complementary and policy makers are often not aware about their possible interactions ex-ante. The evaluation of climate and energy policy instruments and their interaction is an issue that has not been explored systematically because of its distinctive characteristics that must be considered and handled with caution. The high degree of complexity, the multi-objective type of decisions and the tendency to include stakeholders' preferences into the evaluation and decision making process constitute the main features of the problem of evaluating climate and energy policies and their interactions. Therefore an ex-ante evaluation framework to assess the degree of interactions between different energy and climate policy interactions needs to be designed and developed. More particular, the use of an evaluation framework that can adequately deal with these characteristics, structure and analyze decision problems at hand, and assist stakeholders to construct their preferences, is indispensable.

Multi Criteria Analysis (MCA) represents an evaluation framework able to deal with complex decision problems, to consider multiple objectives and criteria, and to integrate stakeholders' preferences into the decision making process in a systematic and constructive manner. MCA describes a structured approach used to determine overall preferences among alternative (policy) options, where the options achieve variant objectives or evaluation criteria. In MCA, desirable objectives are specified and corresponding criteria are identified. MCA evaluation techniques have the ability to consider different types of criteria (e.g. economic, environmental, and social), which are relevant to the decision problem at hand. In addition MCA is capable to process different type of information by incorporating both objective (policy options' performances) and subjective (stakeholders' preferences) information and data.

In energy and climate policy evaluations, stakeholders and decision makers implicitly or explicitly express their perceived relative importance between criteria by assigning weighting factors to them. Furthermore, stakeholders' objectives and policy priorities should be taken into account and even get incorporated into the decision making process in a structured and transparent way. This process can render decisions more defensible and acceptable. There are several weighting methods applied to elicit weighting preferences from relevant stakeholders and decision makers for energy and climate policy evaluation allowing them also to participate in the decision making process (Hobbs and Horn, 1997, Bell et all, 2001, 2003)

The significance of participatory methods in policy making decisions, especially for climate change and energy issues, is gradually acknowledged in the policy world. In IPCC (2001) Assessment report, the usefulness of participatory methods in the decision making processes was exclusively mentioned. An inventory of such tools revealed that there was no consistent method, widely accepted that could assess their individual processes. Different country-specific social characteristics favour different methods. Nevertheless, an important recognition is that participatory methods can 'inform the public, incorporate public values, assumptions and preferences into decision making…and reduce conflict among stakeholders'. The fourth assessment report of IPCC (2007) refers to participatory methods as means of selecting preferred mitigation options and sets such models next to traditional cost-benefit analysis and green accounting. These methods belong to the 'institutional framework' where measures transforming the institutional structure are among the most accepted ones for decoupling development with energy use and emissions. In comparison to Cost Benefit Analysis method, Multi Criteria Analysis is argued that should be preferred in evaluation cases where there are criteria that cannot be easily accommodated in CBA and when there are benefits that cannot be quantified and valued (UNFCC, 2002). In addition, UNEP in an attempt to provide guidelines and analytical frameworks for consistent development of national GHG gases abatement scenarios suggests a combined framework of top – down and bottom - up approaches were explicitly stresses the important role of Multi Criteria Analysis for the selection of national technical options for GHG gases mitigation (UNEP, 1994).

MCA approaches in Energy and Climate policies evaluation

MCA evaluation approaches have been applied widely the last two decades for energy policy evaluations and have been explored in various energy related decision making contexts (Hobbs and Meijer, 2000). In particular MCA approaches have been used for incorporating public values in energy future scenarios evaluation (Keeney et al, 1990), for evaluation of alternative energy plans (Hobbs and Horn, 1997), for elicitation of stakeholders' values for indirect valuation of energy externalities (Diakoulaki and Grafakos, 2004), for designing and evaluating integrated energy and environmental policies (Greening and Bernow, 2004), for integrated assessment of energy analysis (Giampetro et al., 2006).

The last years the increasing use of MCA methods for climate policy evaluation indicates a trend towards more integrative and participatory multi – objective approaches (Bell et al., 2001, 2003, Brown and Corbera, 2003, Borges and Villavicencio, 2004, Diakoulaki et al., 2007, Konidari and Mavrakis, 2007).

Bell et al. (2001) organised a workshop with climate policy experts and policy makers and explored thoroughly the applicability and the usefulness of different MCA methods in integrated assessment and evaluation of climate policies. Participants were in favour of the possibility to revise their initial preferences on weights either by using one method that allowed revision or by combining different weighting techniques. The weighting methods examined at this study were point allocation, swing, analytical hierarchy process, and trade off. One of the main conclusions of the study was that the use of multiple methods can enhance understanding of the policy problem and trade offs between criteria (Hobbs and Horn, 1997, Bell et all, 2001).

Bell et al. 2003 tested how MCA techniques can assist users to incorporate their background knowledge, to improve understanding of trade offs, and to perceive importance of value judgements by ranking hypothetical greenhouse gas mitigation policies. In particular, a weighting method that was mostly recommended for actual decision making was the reconciliation of weighting factors from multiple methods (Bell, et all., 2003).

Brown and Corbera (2003) examined the implications of forest carbon projects for different aspects of equity and development by applying a participatory stakeholder Multi-criteria Analysis to explore the range of stakeholders, their roles, interests and perspectives, for a case study in Mexico. They constructed a simple multi-criteria exercise consisting of the interviewee's qualitative evaluation of a set of sustainable development indicators for project assessment and monitoring, which reflected the carbon, ecological, and social dimensions of forestry carbon projects. The interviewees were asked to value the criteria using a set of qualitative techniques (ranking, qualitative scales and percentage weighting) in order to map their interests and expectations.

Following UNEP's analytical framework (1994) for developing GHG emissions abatement strategies, Borges and Villavicencio (2004) applied a MCA approach to involve all relevant stakeholders on the evaluation of emission's abatement options in Peru. This approach onwards comprises the base for communication by the Peruvian government to the UN Framework on Climate Change. The combination of scenario validation with multi criteria prioritisation of options provided a structured interactive process by which the two procedures could evolve parallel taking into consideration the different stakeholders' perspectives.

An MCA technique has been also applied for the elicitation of stakeholders' views and trade offs for the assessment of different energy scenarios and more particularly for assessing the role of carbon dioxide capture and storage. The direct point allocation weighting method was used to facilitate respondents to weight the importance of evaluation criteria (Shackley and McLachlan, 2007). The same weighting method (direct point allocation) has been incorporated to a MCA decision support tool developed for the regulation of emissions from international civic aviation sector (Solomon and Hughey, 2007). The authors stress the transparency of the decision making process and the ability of MCA to highlight different perspectives and further state the necessity of development and application of such tools for approaching international environmental problems. A pair – wise approach based on Analytic Hierarchy Process (Saaty, 1980) has been applied to elicit stakeholders and experts' views for the assessment of climate change impacts on agricultural land use (Abildrup et al., 2006).

Based on literature review and conclusions of MCA applications (Bell et al., 2001, Bell et al., 2003), decision making for evaluation of energy and climate policies appears to lack of an integrated multi criteria weighting method that combines abilities of different techniques' to derive stakeholders' preferential information and perspectives in a structured, transparent and informative manner. Nowadays, while use of softwares that combine different methods widely increases, practitioners and analysts do not need to use the methods in puristic disconnected manner and they are able to refine the methods to be more suitable for a particular decision making situation. As Poyonen et al. (2001) clearly stated "the strict boundaries between different methods are already passed history". The scope of this paper is to develop an integrated multi criteria weighting methodology that consists of ranking and weighting parts by integrating their capabilities to derive verbally and arithmetically stakeholders' preferences. This methodology has been developed as a basic component of a decision support tool for Energy and Climate Policy Interactions (ECPI) recently elaborated [2].

The structure of this paper has as follows. In section 2 we review and shortly describe the main multi criteria weight elicitation methods. In section 3, the MCA evaluation approach is explained and the evaluation criteria for climate and energy policy interactions are presented. Section 4 is dedicated for the analytical description of the developed weighting methodology. Evaluation responses gathered during the testing phase of the methodology are discussed at the fifth section. Conclusions and further research prospects are drawn to the final section.

2. Classification of weighting approaches

There are various methods to determine criteria weights, which can be classified in different ways: algebraic or statistical, decomposed or holistic, direct or indirect, and compensatory or non compensatory. Algebraic methods often compute the n weights from a set of n – 1 judgments using a simple system of equations. Statistical procedures use regression analysis based on redundant set judgments. Decomposed procedures are based on the comparison of one or one pair of criteria at a time and holistic methods are based on holistic evaluation of alternatives where decision makers / stakeholders consider both criteria and alternatives during the expression of their preferences. Direct methods require the respondents to compare two criteria in terms of ratio judgments and indirect procedures ask them to express preference judgments to derive criteria weights.

[2] More information on the ECPI tool can be retrieved from the reports on the tool's webpage http://www.rug.nl/edrec

Based on the concept of compensation and trade offs between criteria, methods can be also distinguished between compensatory and non – compensatory. Compensatory weighting techniques are used in Multi Attribute Utility (MAU) methods, while non-compensatory ones are used mainly in outranking methods. The former assume strong compensation (trade offs) between criteria and are used as scaling factors, while the latter reject this assumption and are used as importance coefficients in the respective aggregation formula. A central concept in compensatory methods is the impact range of criteria performance (impact range sensitivity), which should be carefully considered by the respondent in order to express the trade offs between criteria.

2.1 Non - compensatory methods

Non-compensatory weighting methods reflect in principle global values about the relative importance of criteria, and do not pay particular attention to the impact range of the specific decision context. The most broadly used non - compensatory weighting methods are the following:

- *Direct point allocation or fixed point scoring techniques* (Hajkowicz et. all 2000, Poyhonen and Hamalainen, 2001)

Direct point allocation method is very simple and straightforward. The respondent is required to distribute a fixed number of points among the criteria. Usually they are expressed as percentages where 100 points are allocated among the criteria. The attribute with the highest score is the most important one.

- *Ratio or direct importance weighting methods* (procedures) (Fischer, 1995, Weber and Borcherding, 1993)

The direct importance (or ratio) methods involve two stages. Initially the stakeholder is required to rank criteria and then to rate them according to their relevant importance. For example the least important attribute can be assigned with a value of 10 and all the others can be rated as multiplies of 10; alternatively the most important criterion can be assigned a value of 100 and all the others can be expressed proportional to it. Then the obtained weights are usually normalized to sum to one.

- *Pair wise comparison techniques*

Pair wise comparisons involve the comparison of each criterion against another one in pairs. The number of pair wise comparisons that should be performed to have all criteria compared to each other is N = c (c- 1) / 2.

The *Analytic Hierarchy Process (AHP*) (Saaty, 1980, 1995) is the most popular technique of using pair wise comparisons. This method requires the user to rate the importance of each criterion in pairs on a nine-point scale, varying from 1 (equally important) to 9 (extremely more important).

- *Resistance to change technique*

This method based on personal construct theory was adopted from Rogers and Bruen (1998) to estimate the relative importance of environmental criteria. This weighting technique is included in *outranking MCA methods* (e.g. ELECTRE). It has elements from the "swing" method (see next section) but also from pair wise comparison techniques. Each criterion is assumed to have two different poles of performance. Those sides of the poles are the desirable and the undesirable ones. Assuming that all criteria are at the desirable side, the stakeholder is required to compare all criteria between each other in pairs and to choose which one is willing to move from the desirable to the undesirable side having the other constant. The total score of each criterion is obtained from the frequency of its resistance to change. Thus a hierarchy of the criteria is determined.

2.2 Compensatory methods

Compensatory weighting methods aim at showing and stressing the hidden dilemmas behind a number of mutually exclusive options evaluated across multiple criteria for making stakeholders become aware of the potential gains and losses implied by their choice in the specific decision context. Thus, it is meaningful to elicit them by taking into consideration the impact range in each single criterion (Keeney, 1992). In this sense, derived weights have no absolute meaning and do not reflect general values in life but only preferences and priorities in the face of considered alternatives. Most widely used compensatory weighting methods are the following:

- *Swing method* (von Winterfeldt and Edwards 1986)

The Swing weighting method at a first stage constructs two extreme hypothetical scenarios, worst and best, where the former presents the worst performance in all criteria (worst score of the examined alternative options) and the latter the corresponding best performance. The stakeholder is required to choose which one of the criteria to move from worst to best performance assuming that the others remain at their worst performance level. The criterion with the most preferred 'swing' is assigned 100 points. Furthermore the stakeholder is asked to select the second most preferred criterion to move from its worst to its best performance level as the second most desirable improvement and consequently to assign less than 100 points to it. This process is being continued for all remaining criteria which are expressed as percentage of the most preferred swing criterion. By normalizing them, weighting factors are determined.

- *Trade off method* (Keeney and Raiffa, 1976)

This method has very strong theoretical foundation. Trade off weighting method constructs two hypothetical alternatives, which differ in two criteria. The first has the best performance on one criterion and the second has the best performance at the second criterion. The respondent is required to choose one of the two alternatives. The selection reveals his preference on a criterion. The respondent is asked to adjust one of the two criteria performances in order to reach the level of indifference between the two alternatives. This can be achieved either by worsening the chosen alternative at the good outcome or by improving the non-chosen alternative at the bad outcome. It is necessary to have n-1 comparisons of pairs of the hypothetical alternatives in order to calculate the weights of n criteria.

- *SMART* (Edwards, 1997)

The simple multiattribute rating technique *SMART* is a typical representative and most commonly used technique of the direct importance weight methods. SMART is a whole process of rating alternatives through weighting criteria. Stakeholder is first required to rank the importance of criteria from the worst criterion levels to the best levels, and then asked to express ratio of importance estimates of each criterion to all others and therefore at the end his judgments can be easily translated into normalized weights. Some weakness of SMART was corrected by the elaboration of SMARTS and SMARTER (Edwards and Barron 1994).

- *MACBETH* (Measuring Attractiveness by a Categorical Based Evaluation Technique) (Bana e Costa and Vansnick, 1994).

This technique considers weights as scaling constants and the weight elicitation procedure is a part of the overall technique. This weight elicitation procedure does not assess the weights directly according to their relevant importance, but it takes into consideration the (impact) range of each attribute. The method integrates "swing", (pair wise) and trade – off elements, providing also the necessary consistency tests for coherency of the procedure. The weights correspond to the concept of trade off; requiring the stakeholder / decision maker to answer how much he is willing to give up from one swing of performance of one criterion to achieve an increase of swing of performance of another one. The question that can be answered assuming that all the criteria are at their worst impact levels, is to assess whether the gain with respect to performance of one criterion by moving from worst to best level is greater or less than the corresponding gains for each other criterion.

- *Conjoint (regression methods)*

This method is the typical example of a holistic and indirect procedure and requires stakeholders to rank or rate the different alternatives according to their preferences. At a next step the analyst using some regression statistical analysis can derive the single value functions and the weights for these functions. The regression procedure is the most common used conjoint method. Conjoint methods derive the relative importance of criteria through an indirect and holistic manner. A necessary precondition to perform the conjoint method is the large number of alternatives and criteria to be able to apply the regression analysis. People tend to ignore or misinterpret many criteria, still important ones, when ranking the multi objective alternatives, treating them in an inconsistent way. This is due to the fact that by using this method stakeholders are not involved on the whole process of elicitation of weights without being asked to reflect and reassess their initial preferences (Hobbs and Meier, 2000).

3. MCA approach for evaluation of energy and climate change policy interactions

In order to incorporate stakeholders' weighting preferences into the evaluation of energy and climate change policy interactions methodology and decision making process, a MCA approach has been used. MCA was deemed appropriate to be integrated in the evaluation framework of energy and climate policy interaction (ECPI) tool for the following reasons:

- Multiple instruments and thus multiple combinations of instruments (policy options) for evaluation can be considered and evaluated by MCA;
- The capability of MCA to deal with multiple criteria and objectives often conflicting, as in the case of multiple criteria in climate and energy policy interactions;
- MCA can incorporate different decision/policy makers' preferences and value systems by the application of the appropriate weighting technique;
- Climate policy interaction is a high complex issue where MCA has the ability to deal with complex policy issues by structuring and analyzing them in a transparent way
- MCA can incorporate and combine objective with subjective type of information (expression of judgments and preferences).

3.1 Criteria for evaluating energy and climate policy instruments

The developed evaluation weighting framework is based on a list of criteria that refer to the three dimensional framework, which consists of energy, environment, and society (OECD, 1997; 2001; IPCC, 2001). A literature review reveals a large number of criteria in climate and energy policy analysis and evaluation. Criteria have been distinguished between four general objective categories, which consist of several specific aggregated top-down criteria, based on various studies (Oikonomou and Jepma 2006; 2008; IPCC, 2007). Evaluation criteria originate from different fields of human activity and their respective objectives, namely addressing climate change mitigation, energy effectiveness, socio-economic objectives, and competitiveness and technological objectives. Similar objectives and criteria are present also in other major studies (for instance IPCC, 2007, OECD, 2001).

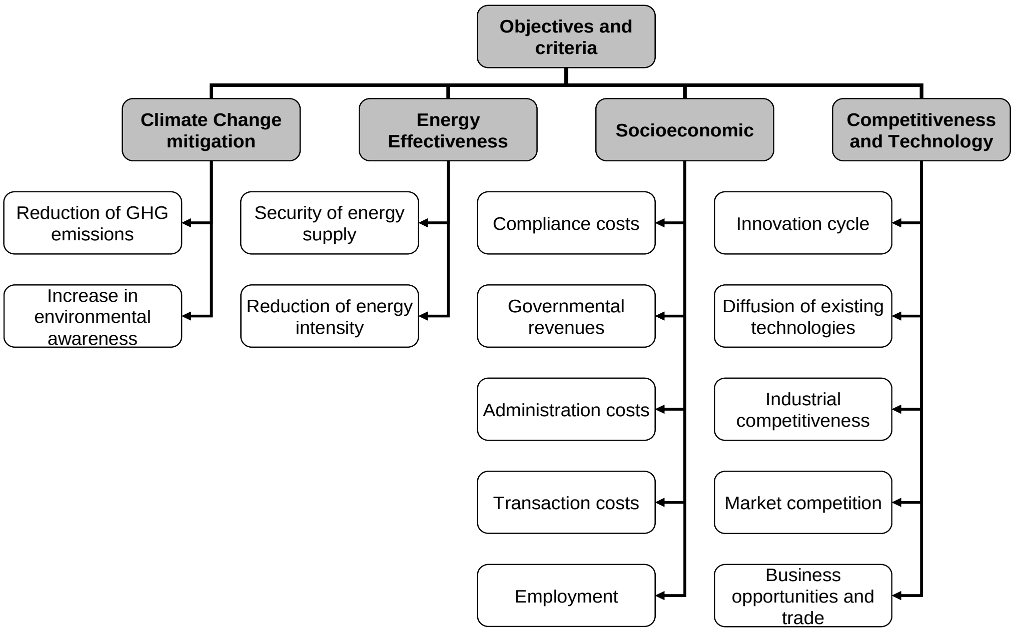
The objectives and criteria are the measures of performance by which the different options of climate policy instruments interaction will be judged. In principle evaluation criteria must fulfil some qualitative attributes as clearly expressed by Belton and Stewart (2002) and Hajkowicz et al. (2000):

- *Operationality* - Being able to specify how well each option of policy interaction meets the objectives expressed by the evaluation criteria.
- *Value relevance* – Linking the concept of each criterion to the final objectives it is meant to represent. In other terms, it presupposes that an objective is comprehensively described by underlying criteria.
- *Preferential independence* – Preferences associated with the performances of each option should be independent of each other from one criterion to the next. This condition has to be met if the sum of weighted averages is to be used to combine preference scores across criteria. Alternatively, preference scores of one option on one criterion should be able to be assigned independently of knowledge of the preferences scores on all other criteria.
- *Decomposability* – Possibility to break down an objective into specific means.
- *Reliability* – A malfunctioning criterion should not render the whole set of criteria unworkable.

- *Measurability* – Degree of measurement of the performance of alternatives against specified criteria.
- *Non-redundancy* – Limiting the number of criteria addressing the same objective, meaning avoidance of duplication of information in criteria.
- *Minimum size* – The number of criteria employed should be only the absolutely necessary to provide representation of policy objectives.

Thus, the selection of evaluation criteria was based on the above conditions and their relevance and link to main four climate and energy policy objectives. A detailed explanation of the selected criteria is present in the ECPI tool's reports. Figure 1 presents the link between objectives and evaluation criteria.

Figure 1: Objectives and selected evaluation criteria



Source: Oikonomou et al. 2008

4. The developed weighting methodology

The design of weighting methodology is based on a decision problem, namely the ex ante evaluation of climate and energy policy instruments interaction. The focus of this study experiment is mainly on the function, the applicability and the interactive elements of the developed weighting methodology with the respondents rather than on a real policy decision problem. Furthermore, the design of the methodology is towards the decrease and minimization of overburden to respondents as has also been stressed from other studies (Bell et al. 2001, Bell et al. 2003). This resulted to the co - application of simple techniques and less data provision to require less time from the stakeholders to spend on the assignment. In order to "force" the users to respond further to the policy problem and more particularly to reconsider, think harder their value system, and deliberate their preferential judgments on evaluation criteria, it is recommended to use parallel multiple techniques than a single one (Hobbs and Horn, 1999, Hobbs and Meier, 2000, Bell et al. 2001). The weighting methodology being presented in this paper integrates two different ways of ranking criteria but also requires respondents to resolve conflicts between them.

The weighting methodology that has been developed to incorporate stakeholders' preferences and derive values for criteria weights is a combination of pair wise comparisons and ratio importance weighting methods, accompanied by a ranking technique for introducing the users to the notion of evaluation criteria but in addition to test the ranking consistency. Weighting techniques that allow respondents to give imprecise rank order information may be a mean to remedy for time consuming, subject to inconsistency weighting techniques and may assist in practical preference elicitation (Hayashi, 2000, Poyonene and Hamalainen, 2001). However, holistic approaches and the judgment of all criteria at once make impossible the consideration of the criteria in a careful and insightful manner. Thus respondents' preference statements cannot be considered as defensible and balanced using only a holistic ranking approach (Hobbs and Meier, 2000).

Therefore, a pair wise criteria approach was applied sequential to the ranking technique to facilitate respondents express their weighting preferences of the criteria in a decomposed way. Since the number of criteria is high, the pair wise comparisons that have been performed (and required are n – 1 in an abbreviated pair wise comparison format. Thus, not all possible pair wise comparisons n (n-1)/2 are presented in the abbreviated format. Pairs are sequentially assigned (as a – b, b – c, c – d, etc.), where the initial criterion and the second criterion in each subsequent pair are randomly assigned, in order to minimize problems with path dependency (Saaty, 1987). A complete ranking of criteria is based on the actual choices and assuming transitive preferences. Keeney (1982) clearly states that one of the basic axioms of Decision Analysis is the transitivity of preferences. If a is preferred to b and if b is preferred to c, then a is preferred to c. Although there is much criticism concerning the transitivity of preferences (Tversky, 1969), the assumption of transitivity is based on the findings of Peterson and Brown (1998) that people are transitive in their preferences revealed through a psychometric method of pair wise comparison method. In this case consistency should be considered guaranteed as long as stakeholders are highly competent and careful, the degree of relevance between the items of the criteria set is the lowest possible, and the value contrast between choices is significant enough (Peterson and Brown, 1998, Strager and Rosenberger, 2006).

The selection and development of this type of pair wise comparisons weighting technique was based on the following reasons:

- The design and application of this weighting technique overcomes the main difficulties that lie in criteria weights elicitation stage and their interpretation which are impact range sensitivity, consistency, hierarchical (splitting) bias (Hayashi, 2000, Poyhonen et al, 2001) and the association of verbal expressions to the 9 points numerical scale:
  a) Given the ex-ante qualitative evaluation of energy and climate policy interactions context, *impact range sensitivity* is not a priority issue to be considered during the current weighting elicitation procedure. Impact range of criteria should be normally presented and stressed during ex post evaluations as necessary and important information. In these cases compensatory

techniques should be used while are characterized by a higher degree of impact range sensitivity. In our study of ex ante evaluation the main objective of the method is to facilitate stakeholders to express their preferences of relative importance of criteria and their global values for certain evaluation criteria. However, being aware of the fact that impact range sensitivity is one of the major issues that should be considered in weights determination, the developed methodology can incorporate impact ranges of criteria. Furthermore the method can provide this type of information to stakeholders while expressing their weighting preferences when a complete set of policy alternatives complements the weighting process. Although the degree of impact range sensitivity cannot be measured, the method can require the stakeholders to be aware about the impact ranges of criteria before expressing their preferences.

b) The division of criteria (or objectives) in value trees and subcriteria categories can either increase or decrease the weight of a criterion (or objective) because respondents cannot adjust their responses enough to a change of a value tree. Furthermore, the variant ways of structuring and then weighting criteria and sub criteria in value trees may also change the rank of criteria, a phenomenon which is called the unadjustment phenomenon or as it is widely used splitting bias (Weber et al., 1988, Poyonen et al., 2001). By developing and applying this weighting technique *hierarchical (and splitting) bias* is avoided as all criteria are compared in pairs without any hierarchical value tree structure and aggregation of criteria at each category.

c) Transitivity of preferences, as discussed in a more detailed manner above, is assumed by applying the abbreviated pair wise comparisons and thus *inconsistency check* is not deemed necessary. Nevertheless, during the development of the weighting method, a simple ranking consistency check has been introduced providing the opportunity to respondents to revise their initial preferences.

d) It was required from stakeholders to express their preferences importance between criteria verbally as an introductive step before asking them to further express their preferences' intensity in a ratio numerical scale which was further accompanied by an automatic visual representation of the selected relative importance between the criteria. By applying this way of verbal expression of preferences and then further asking ratio numerical preferences the methodology avoids the weakness of associating verbal expression to a preset numerical scale like Saaty's 1-9 point numerical scale which has been subject of criticism for this reason.

- The number of criteria renders the application of other weighting techniques impossible and thus restricts the possibility to use only few weighting techniques that can consider a high number of criteria.
- As a user friendly weighting procedure (structured, simple, step by step), it does not require a lot of time and effort from stakeholders and it therefore reduces the cognitive burden required by them.
- It can be used by many and different individuals simultaneously either in a form of individual interview or by electronic communication. In addition it can be used within groups to identify trends, preference differences and conflicts, and rise in depth discussion for the evaluation problem at hand.

- The pair wise comparisons method has its roots at the Analytical Hierarchy Process (Saaty, 1980), a well-known, broadly used method with a strong theoretical foundation.

4.1 Stages of the proposed methodology

The presented method consists of three methodological stages, as described below.

First stage: Ranking of criteria

During the first stage of the weight elicitation process respondents are required to rank criteria according to their preferences, from the most preferred to the least preferred criterion. This stage is an introductory phase to the weighting procedure as is mainly used to familiarize in a simple way the stakeholders to the concept of ranking and comparing the criteria in a holistic approach.

Second stage: Pair wise comparisons

During the second stage respondents are required to express which criterion they prefer at each pair wise comparison. Furthermore, they are asked to express how much more they prefer one criterion against the other according to their level of preferences. There are five levels of intensity of preferences have been defined in verbal expressions. Every level of preference intensity is associated to certain numerical ratio level(s) of a 10 points scale between 0 – 1 (0.1, 0.2, 0.3, etc.):
The five levels of intensity of preference values verbally expressed, associated to 10 levels of numerical preference values, are the following:

- equally preferred (1),
- almost equally preferred (0.9),
- moderately preferred (0.6 – 0.8),
- strongly preferred (0.3 – 0.5) and
- very strongly preferred (0.1 – 0.2).

The user is also required to express the ratio (percentage) of the least preferred criterion comparing to the most preferred according to his level of intensity of preference. The user is assisted by the use of a computer aided excel tool which has been developed for this purpose. The respondent can choose from a 10 points scale from 0 to 1 according to the intensity of his preferences and ratio importance after expressing them initially in a verbal mean as explained above.

Then an inner, thin bar will be moved automatically according to his preferences within the thick bar (graphic representation) of the preferred criterion providing him the visual representation of his preferences of relative importance between the pair of compared criteria. When the respondent expresses the relative importance (relative scores) between criteria at every pair comparison, weighting factor (Wi) of criterion i is obtained by the following formula:

$$Wi = \frac{RSi}{\sum_{n-1} RS}$$

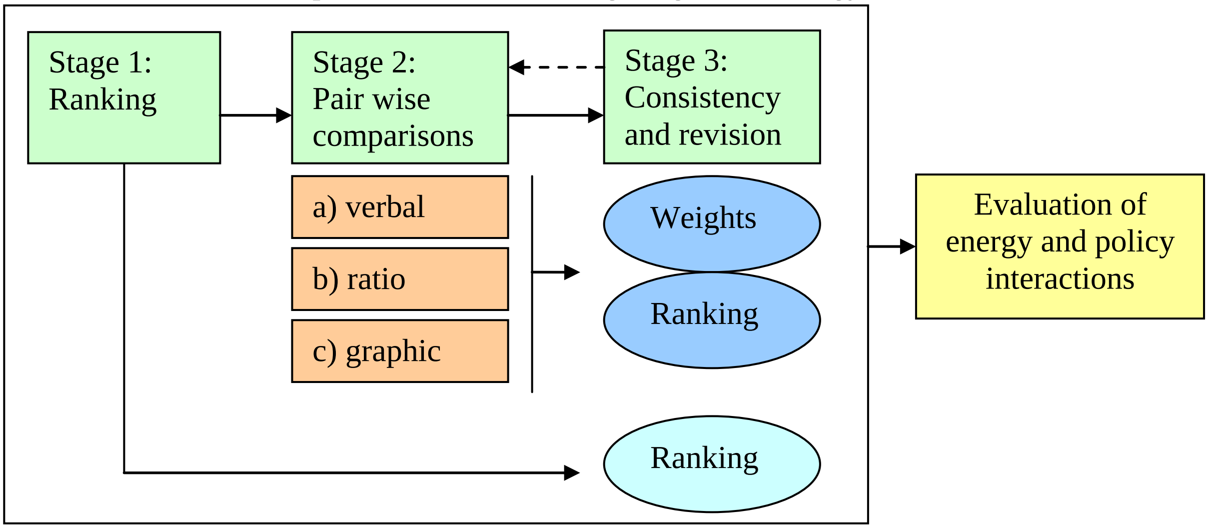
Where RSi is the Relative Score of criterion i compared to the criterion j during the pair wise comparison and Σ (RS) is the sum of Relative Scores of all criteria (n) after completing the whole set of abbreviated pair wise comparisons (n – 1). The first criterion has been assigned with relative score of 1, to be used as basis reference relative score for the calculation of the relative scores of the sequential pair wise comparisons of criteria.

Using the computer aided tool for the calculations of relative scores, weighting factors and ranking of criteria are obtained automatically, while respondents can perceive a visual representation of every pair wise comparison that they perform as described above (see appendices III, IV, V).

Final stage: Possibility for revision and consistency check

The computer aided decision tool is interactive in a sense that stakeholders, after observing the derived weighting factors and ranking of criteria, can revise their preferences by performing again some or all pair wise comparisons accordingly. The obtained ranking of criteria at the second and main stage of the weighting process is compared to the ranking of the first stage of the procedure and check if there are any differences. Thus, the use of the first stage of the process is not only to introduce the decision makers to the notion of criteria importance but also to provide a simple consistency ranking check. Then the respondent can decide which ranking better reflects his preferences and thus he can revise his choices accordingly. If there are no differences of his preferences between the rankings of the two stages, then the respondent is consistent. If the rankings are different and the respondent thinks that the ranking of second stage of weighting process reflects better the respondent's preferences then the process has been completed. If the stakeholder believes that the first selected ranking represents better the preferred ranking importance of criteria then he is required to revise his preferences during the pair wise comparisons (second stage) accordingly. Normally, during the pair wise comparisons stakeholders are "forced"/facilitated to better think about their preferences of relative importance between evaluation criteria following a structured, decomposed and systematic way.

Scheme 1: Schematic representation of the weighting methodology



5. Evaluation responses

The majority of users showed more confidence to the ranking resulted by the pairwise comparisons and declared that the resulted weights expressed sufficiently their preferences. It should be also noticed that more experienced users chose the pairwise comparisons method as their intuitively more preferred ranking method between these two; this fact is in line with the principle that more complex methods demand more experienced users to be fully exploited.

5.1 Insights and recommendations

The method is tested by a team of individual stakeholders and experts working on climate and energy issues in academic, governmental, research and energy institutions and organizations. The following conclusions were synthesized from their responds and comments written on questionnaires offered after the application of the combined methodology of criteria weighting method (pairwise comparisons) and criteria ranking method.

Co-application of both methods
All participants acknowledged that the combination of two methods and their different level of application – since the pairwise comparison is a weighting method as well– were practical because it introduced an introductory session (the spontaneous ranking step). Being free to rank the criteria in a holistic way without an immediate obligation in expressing their relative weights, users adapted gradually to the problem and the more accurate expression of their preferences through the pairwise comparisons was facilitated.

Correlation between the expression of experts' preferences and the available policy alternatives
A lot of users expressed their desire to address a defined set of policy alternatives before expressing their preferences, on the grounds that their preferences could be modified according to available alternatives. This comment implies the weakness of the less experienced experts to fully recognise the posed problem, which is no other than the climate and energy policy formulation. However, users may gain some more understanding and feel more confidence in the case that a presentation of the policy selection problem and of the alternative policies set precedes the application of the tool. It should be mentioned here that the intension was to test the methodology for deriving stakeholders' relative importance factors reflecting their general values towards the criteria. It is acknowledged that this is a simplification of a real policy problem which constitutes a future research direction.

"Coaching" of users
The testing of the tool is conducted remotely by the respondents and the communication link was limited to e-mail communications. Although the remote testing added to the independence and the creditability of the results, the limited communication between the method developers and the tool users had the drawback of missing a more appropriate coaching session for the less familiar with preference elicitation techniques stakeholders. These users thus faced some difficulties to use this method in the most efficient way and some of their comments are not in line with the

comments of the more experienced ones. In retrospect, a more extensive testing session of this method should incorporate a more thorough cooperation between developers and users (also during a stakeholders' workshop), including some joint test runs of the process and the potential for interaction during the test phase.

Scaling of preference
Most participants welcomed the ability of the pairwise comparisons method to integrate, beyond the verbal expression of preference, a quantitative and visual expression. On the other hand, some of them remarked that the use of a quantitative scale, which decomposes the verbal values to a more analytic set of values, might provide a deceptive sense of accuracy. This remark is based on the fact that people in common are accustomed to use a limited verbal set of preference expressions. Taking into account this remark, the application of the method with only an univocal quantitative expression of the verbal expressions set can be explored in the future.

Preference independence
Some users implied that expressing their preferences was encumbered by overlaps between the selected criteria. This notice highlighted two critical points. The first point is the need for a more detailed presentation of each criterion in order to minimise the cognitive burden for the respondents to fully understand the explicit significance of each criterion. The second is the need to encourage them to consider the possibility of preference independence; even where they may still feel that there is some physical dependence. However, this requires the willingness and the flexibility of the users to accept such an approach, and such requirements are not always the easiest to be fulfilled. It was recognised also that some users feel more comfortable to express their preferences in a holistic way where others can better express them in a more structured decomposed manner.

6. Summary and conclusions

The ex-ante evaluation of climate change and energy policy instruments interaction is a subject that demands an approach that should be as simple as possible, in order not to add excessive burden to stakeholders. An important component of this approach is the integration of the stakeholders' weighting preferences on energy and climate change policy evaluation criteria, in a simple, as well as transparent and systematic way. In order to derive stakeholders / decision makers' preferences for criteria weights, an integrative constructive weighting methodology has been developed and applied, which is a combination of pair wise comparisons and ratio importance weighting methods. This weighting method introduces stakeholders to a "warming up" holistic approach for ranking evaluation criteria and then requires them to express their ratio relative importance in pair wise comparisons of criteria by providing them an interactive mean with verbal, arithmetic and visual representation of their preferences. The excel aided tool, elaborated for this purpose, facilitates stakeholders to intervene and participate in the evaluation and decision making process and set their preferences that determine the outcome of policy interactions evaluation.

Stakeholders and experts in the climate policy field have tested this application and their evaluation views on the application of the methodology have been stated and commented in this paper. They provided their approval for the co-application of these methods, since it permits the gradual approach to the evaluation problem and the more accurate expression of stakeholders' preferences. In addition, stakeholders provided feedback concerning the correlation between the level of experience and relevant knowledge of the users, and the potential of this methodology to obtain the real preferences of the stakeholders.

In this context, there are future research prospects concerning the enhancement of this methodology. The first concerns the need to explore how the quantitative and verbal assignment of preference values can become more representative and simultaneously accepted by the stakeholders. Another critical prospect involves the encounter with the weakness to cover fully independent criteria and the facilitation of the expression of some preference even in the case that some of the criteria have some physical interdependence. In addition, it should be recognised that further research merits on the interrelations between the stakeholders' preferences in a set of proposed alternatives in a real evaluation problem of climate change and energy policy interactions where performances of different options and hence their impact range are also taken into account. Moreover there is much space for testing the contribution of the methodology during participatory stakeholders' workshops where its use could assist them by facilitating in depth discussion, valuable for more informative and transparent decision making process and more defensible policy decisions.

Acknowledgments

The authors are grateful to Professor Benjamin Hobbs from Johns Hopkins University (Baltimore) for his fruitful comments on the developed integrated weighting tool.

References

Abildtrup, J., Audsley, E., Fekete-Farkas, M., Giupponi, C., Gylling, M., Rosato, P. and Rounsevell, M. (2006). Socio-economic scenario development for the assessment of climate change impacts on agricultural land use: a pairwise comparison approach. *Environmental science and policy,* 9, 101-115.

Bana e Costa, C. and Vansnick, J.C. (1994). MACBETH: An interactive path towards the construction of cardinal value functions. *International Transactions in Operational Research,* 1 (4), 489-500.

Bell, M., Hobbs, B., Elliott, E., Ellis, H. and Robinson, Z. (2001). An evaluation of Multi-criteria methods in an integrated assessment of climate policy. *Journal of Multi-Criteria decision analysis,* 10, 229-256.

Bell, M., Hobbs, B. and Ellis, H. (2003). The use of multi-criteria decision-making methods in the integrated assessment of climate change: implications for IA practitioners. *Socio-Economic planning sciences*. 37, 289-316.

Borges and Villavicencio, (2004), Avoiding academic and decorative planning in GHG emission abatement studies with MCDA: The Peruvian case, *European Journal of Operational Research*, 152, 641 - 654

Diakoulaki, D., Georgiou, P., Tourkolias, C., Georgopoulou, E., Lalas, D., Mirasgedis, S. and Sarafidis, Y. (2007). A multicriteria approach to identify investment opportunities for the exploitation of the clean development mechanism, *Energy Policy*, 35, 1088-1099.

Diakoulaki, D., and Grafakos, S., (2004), *Multicriteria Analysis*, In: European Commission, ed. ExternE - Externalities of Energy: Extension of accounting framework and Policy Applications. Luxembourg-Ville, Luxembourg: Directorate - General for Research, Sustainable Energy Systems

Fischer, G., (1995), Range sensitivity of attribute weights in multiattribute value models, *Organizational Behavior and Human Decision Processes*, 62, (3), 252 – 266

Giampetro, M., Mayumi, K., and Munda, G., (2006), Integrated assessment and energy analysis: Quality assurance in multi – criteria analysis of sustainability, *Energy*, 31, 59 – 86

Greening, L., and Bernow, S., (2004), Design of coordinated energy and environmental policies: use of multi-criteria decision-making, *Energy Policy*,32, 721 – 735

Hajkowicz, S., Young, M., Wheeler, S., MacDonald, D. H. and Young, D. (2000). *Supporting Decisions, Understanding Natural Resource Management Assessment Techniques*. A report to the Land and Water Resources Research and Development Corporation. Australia, CSIRO Land and Water.

Hayashi k., 2000, Multi criteria analysis for agricultural resource management: A critical survey and future perspectives, *European Journal of Operational Research*, 122, 486 - 500

Hobbs B. and Horn G., Building public confidence in energy planning: a multimethod MCDM approach to demand-side planning at BC gas, *Energy Policy*, 25, (3), 357–375, 1997.

Hobbs, B. and Meier, P., (2000), *Energy Decisions and the Environment: A guide to the use of Multicriteria Methods*, Kluwer Academic Publishers, Dordrecht

IPCC (2001). Climate Change 2001: Mitigation, Contribution of Working Group III to the third assessment report of the Intergovernmental Panel of Climate Change.

IPCC (2007). Climate Change 2007, Mitigation, Contribution of Working Group III to the fourth assessment report of the Intergovernmental Panel on Climate Change.

Keeney, R., (1992), *Value - Focused Thinking: A path to Creative Decision Making*, Harvard University Press, Cambridge, Massachusetts

Keeney, R., Von Winterfeldt, D, and Eppel, T., 1990, Eliciting public values for complex policy decisions, *Management Science*, vol. 36, No9, 1011 - 1030

Keeney, R. and Raiffa, H., (1976), *Decisions with Multiple Objectives*, J. Wiley & Sons, New York

OECD (1997). Evaluating economic instruments for environmental policy. (ed. Smith and Vos). Paris, OECD.

OECD (2001). Environmentally related taxes in OECD countries: issues and strategies. Paris, OECD.

Oikonomou, V. and C. Jepma (2008). "A framework on interactions of climate and energy policy instruments." *Mitigation and Adaptation Strategies for Global Change* 13(2): 131-156.

Oikonomou, V. and C. J. Jepma (2006). A Qualitative Methodology on Analyzing Climate Change and Energy Policy Instruments. Groningen, the Netherlands, University of Groningen, Energy Delta Research Centre (EDReC).

Peterson and Brown (1998), Economic Valuation by the Method of Paired Comparison, with Emphasis on Evaluation of the Transitivity Axiom, *Land Economics*, 74, (2) 240 - 261

Poyhonen, M. and Hamalainen, R., (2001), On the convergence of multiattribute weighting methods, *European Journal of Operational Research*, 129, 569 – 585

Poyonen, M., Vrolijk, H., and Hamalainen, P., (2001), Behavioural and procedural consequences of structural variation in value trees, *European Journal of Operational Research*, 134, 216 – 227

Rogers, M. and Bruen, M. (1998). A new system for weighting environmental criteria for use within ELECTRE III. *European Journal of Operational Research*, 107 (3), 552-563.

Saaty, T.L. (1980). The Analytic Hierarchy Process. McGraw-Hill. New York.

Saaty, T.L. (1987). Concepts, theory, and techniques rank generation, preservation and reversal in the analytic hierarchy decision process. *Decision sciences*, 18, 157 – 177.

Shackley, S. and McLahlan, C. (2006). Trade-offs in assessing different energy futures: a regional multi-criteria assessment of the role of carbon dioxide capture and storage. *Environmental Science and Policy*, 9, 376-391

Solomon, D. and Hughey, K. (2007). A proposed Multi Criteria Analysis decision support tool for international environment policy issues: a pilot application to emissions control in the international aviation sector. *Environmental Science and Policy,* 10, 643-653

Strager, M. and Rosenberger, R. (2006). Incorporating stakeholder preferences for land conservation: weights and measures in spatial MCA. *Ecological Economics*, 58, 79-92.

Tversky, A. **(**1969). Intransitivity of Preferences. *Psychology Review*, 76, 31-38.

UNEP, (1994), UNEP Greenhouse Gas Abatement Costing studies – Phase two. UNEP collaborating centre on Energy and Environment, Riso, Denmark

UNFCC, (2002), Input of the least developed countries expert group on the improvement of the guidelines for the preparation of national adaptation programmes of action, FCCC/SBI/2002/INF.14, United Nations Office Geneva, Switzerland

Von Winterfeldt, D. and Edwards, W. (1986). Decision analysis and Behavioural research. Cambridge University Press, New York

Weber, M. and Borcherding, K., Behavioral influences on weight judgments in multiattribute decision making, *European Journal of Operational Research*, 67, 1 – 12

Weber, M., Eisenfuhr F., and Von Winterfeldt, D., (1988), The effects of splitting attributes on weights in multiattribute utility measurement, *Management Science*, 34, (4), 431 - 445

Appendix I: Instructions provided within the excel tool

The weighting methodology - Pair wise Comparisons (abbreviated)
Instructions to fill in the excel sheet

1. Click on the first worksheet (Description of criteria) and read the brief description of criteria
2. Move to the second worksheet (1st STEP-CRITERIA RANKING)
3. Rank the 14 criteria in descending order according to your preferences. Attention, choose a criterion only from the dropdown list of the **green** cell.
4. Once you complete the ranking, move to the next worksheet (2nd STEP-PAIRWISE COMPARISONS)
5. For each of the 14 pairs of criteria, fill the 3 green cell in the following order:
    a. Select the criterion you **prefer** between the 2 criteria, from the dropdown list of the **green** cell.
    b. Indicate **verbally** the level of your preference, from the dropdown list of the **green** cell.
    c. Indicate **numerically** the level of your preference, from the dropdown list of the **green** cell.
6. Once you complete the comparisons, move to the worksheet TOTAL RESULTS-EVALUATION
7. If you think that ranking and weights are **not** representatives of your actual preferences then go back to step 2 and modify the pair wise comparisons according to your actual preferences, otherwise move downwards and answer the 4 simple **questions** for evaluating the method

Appendix II: Step 1 – Ranking of criteria

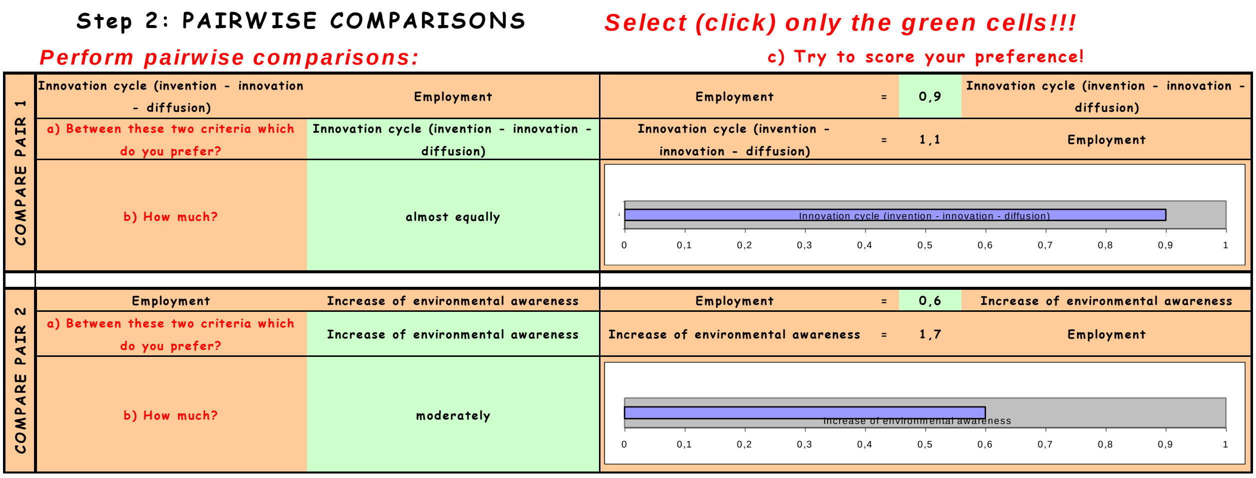
Step 1: SPONTANEOUS RANKING

Select only the green cells!

List of Criteria	Rank	*Rank the criteria according to your preferences*
Innovation cycle (invention - innovation - diffusion)	1	
Employment	2	
Increase of environmental awareness	3	
Market competition	4	
Compliance cost	5	
Governmental revenues	6	
Administration costs	7	
Transaction costs (search, information, negotiation, approval, monitoring, insurance)	8	
Diffusion of existing technologies	9	
Security of supply	10	
Reduction of GHG emissions	11	
Competitiveness	12	
Reduction Energy intensity	13	
Business opportunities and trade	14	

Appendix III:
Step 2 - Pair wise comparisons (abbreviated) and example in excel tool

- **Innovation cycle vs Employment**
- **Employment vs Increase of environmental awareness**
- **Increase of environmental awareness vs Market competition**
- **Market competition vs Compliance cost**
- **Compliance cost vs Security of supply**
- **Security of supply vs Governmental revenues**
- **Governmental revenues vs Reduction of GHG emissions**
- **Reduction of GHG emissions vs Administration costs**
- **Administration costs vs Competitiveness**
- **Competitiveness vs Transaction costs**
- **Transaction costs vs Reduction Energy intensity**
- **Reduction Energy intensity vs Diffusion of existing technologies**
- **Diffusion of existing technologies vs Business opportunities and trade**



Appendix IV: Weighting and ranking results (presented in excel tool)

Criteria		Weighting Factors	RANK
Innovation cycle (invention - innovation - diffusion)		0.13	2
Employment		0.08	5
Increase of environmental awareness		0.16	1
Market competition		0.05	9
Compliance cost		0.05	8
Governmental revenues		0.02	14
Administration costs		0.04	11
Transaction costs (search, information, negotiation, approval, monitoring, insurance)		0.05	10
Diffusion of existing technologies		0.08	6
Security of supply		0.03	12
Reduction of GHG emissions		0.11	3
Competitiveness		0.06	7
Reduction Energy intensity		0.11	4
Business opportunities and trade		0.03	13
Total		1.00	

	Rank
ok	Increase of environmental awareness
ok	Innovation cycle (invention - innovation - diffusion)
ok	Reduction of GHG emissions
ok	Reduction Energy intensity
ok	Employment
ok	Diffusion of existing technologies
ok	Competitiveness
ok	Compliance cost
ok	Market competition
ok	information, negotiation, approval, monitoring,
ok	Administration costs
ok	Security of supply
! SAME WEIGHT WITH THE PREVIOUS ONE	Business opportunities and trade
ok	Governmental revenues

Appendix V: Graphic representation of weighting results (excel tool)