Meta-analysis for comparative environmental case studies: methodological issues

Benedetto Matarazzo
Faculty of Economics, University of Catania, Corso Italia, Catania, Italy
and
Peter Nijkamp
Department of Spatial Economics, Free University, Amsterdam, The Netherlands

Meta-analysis: introduction and scope

Meta-analysis is increasingly recognized as a potentially important analytical framework for comparative research that aims to draw inferences on common issues with different but allied empirical backgrounds (Hedges and Olkin, 1985). The purpose of meta-analysis is to combine findings from separate but largely similar studies (in terms of subjects, hypotheses, phenomena, etc.). Such studies may be suitable for the application of a variety of analysis techniques (common literature review, formal statistical approaches, etc.) for combining, comparing, selecting or seeking out common elements, relevant results, cumulative properties etc. from a broad set of individual cases (see Button and Jongma, 1995).

The aim of this paper is to illustrate and tackle some fundamental questions related to the specific methodological complexities inherent in meta-analysis (techniques to be adopted, selection of case studies, etc.) with particular reference to syntheses in the field of environmental economics. For this purpose we identify six different but interconnected levels of analysis, which, in our opinion, are particularly important both from a purely methodological point of view, and for operative and interpretative reasons. In the following sections we will analyse the specific problems and objectives related to each of these stages by underlining their most relevant methodological aspects, also in relation to the principal objectives of the analytical synthesis.

The principal aims of meta-analysis may be summarized as follows:

- to summarize relationships, indicators etc. in policy-oriented studies;
- to compare, evaluate and rank different studies on the basis of a series of relevant criteria;
- to average (and weight) estimated values, parameters etc. found in different studies;
- to identify common background elements in such studies;

• to aggregate studies by considering complementary results or perspectives;
• to compare, evaluate and rank different methods or alternative policy choices applied to the same (or related) questions;
• to consider and interpret factors (moderator variables) which may be responsible for different results achieved in rather similar studies;
• to correlate the aggregated data of each study with other characteristics of that study.

It goes without saying that meta-analysis is at first glance a fascinating analytical tool. It should be noted, however, that this type of approach is characterized by considerable methodological complexity, due not only to the specific objectives which are set in each case, but also - and above all - to the intrinsic nature of the studies themselves, which prove to be markedly “transversal”, both horizontally and vertically. Transversality refers to both the intrinsic heterogeneous nature of the various studies and to the various different empirical or political processes addressed in these studies. It is therefore not only necessary to make a suitable identification and selection of similar studies and to analyse them by means of an appropriate research technique (horizontal transversality), but also to adopt a vertical orientation, from the stage of identifying the specific problem to be studied to that of using the results of the study in an operative way or in a policy context.

In light of the above observations, it is therefore possible to identify the following different analysis “levels”, each of which assumes a particular importance from a methodological point of view:

• Real-world level: this indicates the space-time reality which is the context of all problems and phenomena studied, which through their interactions constitute a single, but complex system with many actors and issues involved.

• Study level: this consists of the identification, definition and description of the problem selected. In general, this involves the formulation, explicit or implicit, of suitable theoretical hypotheses regarding the phenomenon studied; the verification of these hypotheses by means of the introduction of a pre-selected model (simplified representation of reality); the use of suitable techniques; the collection and/or elaboration of particular data; and the presentation - and usually also the interpretation and the analysis - of the results of this study.

• Pre-meta-analysis level: this consists of defining, explicitly and accurately, the object and the objectives of the synthesis to be carried out, and of indicating in particular the specific problems to be solved and the dimensions established in terms of time and space. The methods and techniques considered most suitable for carrying out successfully the planned research are also defined at this stage.
Study selection level: once the object and objectives of the synthesis of a certain issue to be carried out have been established, it is necessary to determine and select – in terms of both quantity and quality – the individual studies to be reviewed, bearing in mind the meta-analytical techniques which are to be adopted and the ultimate aims of the synthesis itself.

Meta-analysis level: at this level of comparative analysis the studies selected are analysed thoroughly and critically by means of a chosen formal technique; also the consistency of the results thus obtained is carefully evaluated, so as to offer a solution to the problem discussed, of which a proper synthesis is presented next.

Implementation level: this constitutes a post-meta-analysis phase, a kind of “feed-back” or application to the real-world, which considers not only the “expected” results obtained by the synthetic study but also – and above all – the effects of the experience acquired; this phase supplies useful indications and practical suggestions (relevant to the problem studied but also to new studies to be carried out, to strategies to be adopted, etc.) which are not necessarily directly or closely connected with the original objectives of the analysis.

The previous remarks show that the steps prior to and during the application of meta-analysis have to be carefully judged and implemented.

Methodological complexity
It is conceivable from the above observations that meta-analysis in general is characterized by a range of peculiar methodological complexities. These difficulties, moreover, become even more noticeable in the passage from applications in the field of natural sciences to those in the field of social sciences in general, and environmental economics in particular. We will now successively discuss the six different levels distinguished above.

The real-world level
With reference to the real-world level, it is noteworthy that, while many natural phenomena are characterized by a certain degree of “regularity” and by some cause-effect relationship, the same is not always true for the field of social phenomena. This is a field in which the cause-effect relationship is often marked by much uncertainty (Van den Bergh et al., 1995b), since social phenomena do not depend solely on the “nature” of things but also on the human behavioural context (limited rationality, subjective preference, mutual influences, rigidities, etc.) and on the adoption of specific policies.

It is therefore easy to observe different characteristics of the same phenomenon or of the same problem, that is, non-random deviations of the data observed; these may depend on the varying times and places of observation; they may be a consequence of the presence (in different degrees even) of a number of conditioning factors (the so-called moderator variables), or of the
different dimensions of the study carried out; or, finally, of the varying span of
time required before the effects of certain social phenomena are felt. This is why
for the social sciences – unlike the natural sciences – it is impossible to speak of
“strict experiments”, and why it is necessary to use the terms “quasi-
experiments” and “quasi-scientific methods” (see Button and Jongma 1995, and
Van den Bergh et al., 1995a).

For these reasons, it is from a technical point of view not always possible or
methodologically correct to carry out statistical analyses. The presence of a
very limited number of stable laws (so-called statistical regularity) and the
difficulties involved in generation of samples which are random, independent
and of equal dimensions, make it in fact often impossible to make a correct and
reliable use of descriptive statistics, of statistical, classical and Bayesian
inference, or of simulation and other techniques. It is often necessary to resort
to modifications of the formulae employed and to the use of specific procedures
which in any case present serious limitations.

The study level
At the individual study level the problem defined a priori is tackled through the
application of the model, method and technique chosen to the sample under
study. With particular reference to the social sciences, therefore, it is plausible
that the effective formulation of the problem is contingent on the data collected
(or actually available). If such data are incomplete, inadequate or insufficiently
homogeneous, it may in fact be necessary to reconsider the entire research
project, including the definition itself of the problem faced. Moreover, the
information available is often not of a quantitative nature and therefore not
certain; this too, leads to further and considerable methodological
complications for the research project at hand (Van den Bergh et al., 1995b)
because of the impossibility of making precise measurements.

The choice of the method to be adopted is clearly linked to the representation
of the system and to the model selected (conceptual bases, hypotheses, data
collection, etc.). With particular reference to environmental system analysis,
there exists here already a fundamental type of uncertainty, owing to the
limitation of “description-reality” (Van den Bergh et al., 1995b). In all sciences,
in general, theories or models are adopted, which may be interpreted as
intellectually built artificial objects (see Ravetz, 1971) in order to outline and
represent selected aspects or dimensions of reality. Obtaining meaningful
results depends therefore on the skill used to build these models. By means of a
suitable application of these, thorough scientific research may then identify new
properties and verify the extent to which they reflect reality.

This process may therefore be imagined in a circular form; starting from
reality, where the phenomena (facts) exist, and going on by induction to the
world of the intellect, where theories and models are formulated from which, by
deduction, inferences may be drawn which, by means of validation, are then
compared with real phenomena. It is therefore possible to observe (Miser, 1993)
that, while the facts – as a result of careful and selected observations of reality
should prove completely objective, the induction, which transforms these facts (raw data) into ideas (information), is a highly personal process, internal and value-loaded. The building of theories and models, moreover, makes use of craft skills and imagination to combine the facts with complementary knowledge (complementary facts, well-established theories or models, etc.). The process of deduction makes use of logical-formal instruments of a mathematical type, in the widest sense of the word; the consequences, generated by deduction from the theories and models introduced, may be checked against empirical facts from reality.

If we accept this concept of science in general, any theory or model is an approximation of a specific and selected portion of reality. This is true for all sciences, to the degree of the portions of reality that they undertake to approximate and summarize in their theory (Kemeny, 1959). This is consequently also - and particularly - true for environmental economics.

Furthermore, uncertainty is to be found at the model level focusing on empirical case studies, where it is necessary to take basic decisions concerning the variables to be taken into account in the model itself: which of these should be taken into consideration, and which should be considered as indicators, which are exogenous or endogenous? (Van den Bergh et al., 1995b). There are also other difficulties inherent in the formulation of suitable indicators, policies and instruments for the case study approach.

In connection with the above, there are finally also problems of computation (techniques, nature of the data, etc.), presentation, interpretation and discussion of the results of individual studies, which naturally may interact and create feedback mechanisms.

The pre-meta-analysis level
In this phase the methodological choices faced are critical in terms of their share, entity and difficulty, as at this stage they determine the progress and the quality of the synthesis in a comparative study to be carried out. This phase constitutes in fact the first effective part of the synthesis, in which it is necessary to come to grips, among other things, with a series of complex prejudicial problems of fundamental importance in order to guarantee a proper quality level for the synthesis aimed at.

Substantially, this phase consists of describing and defining - explicitly, accurately and with the greatest possible precision - the object and the peculiar objectives of the synthesis to be carried out, indicating, in particular, the specific problems to be tackled and the dimensions established in terms of time and space. Subsequently and in accordance with this, the methodologies and techniques considered most suitable for a correct completion of the programme are then to be determined.

One implicit objective of a general nature in all syntheses is that of reducing the level of subjectivity which is inevitably present in each individual study. This is particularly true in the field of the social sciences and hence in environmental economics, where there is a great variety of particular problems.
caused by the use of an enormous diversity of different methodologies – generally not standardized – in the individual case studies; by the varying objectivity with which the information has been collected; by the existence of studies which are not always well-designed and which do not always deal with appropriate questions; and by a frequent difference in output measures and methodologies for summarizing data (see Button and Jongma, 1995).

In this phase we consider it therefore particularly useful to establish a correct methodological approach, and in particular a suitable relationship between the principal objectives of the synthesis and the chosen techniques. The principal objectives of meta-analysis have already been mentioned above. To reach these results it is possible to adopt a variety of techniques (Van den Bergh et al., 1995a). It is clear that the choice of the meta-analytical techniques to be used proves to be closely connected with the objectives of the synthesis. For example, if the principal object of the synthesis consists of reviewing and grouping a number of different studies characterized by a low level of homogeneity in order to identify the common traits, similar behaviour (spontaneous or induced) or complementary results, and to point out those fields and phenomena not yet examined in sufficient detail (and therefore worthy of further, more intensive, study), traditional approaches, e.g. of a qualitative type, may be usefully, – though perhaps not exclusively – employed.

If the purpose of the synthesis is to identify relationships between the variables studied, evolutionary tendencies, synthetic indicators of calculated values or estimated parameters (or other common elements which can be effectively described in quantitative terms), the use of suitable statistical techniques may be particularly appropriate.

If one of the aims of the analysis is a qualitative confrontation, a comparative evaluation, a ranking of individual studies (also in light of the particular methodologies and techniques used in each of these) and if the results of each individual study can be interpreted as “criteria”, then the use of quantitative or qualitative multi-criteria techniques seems most suitable. If the single studies are to be grouped and classified according to some technical characteristics, data, results or policy options – in the presence of more than one (even qualitative) “attribute” – rough set analysis is likely to be the most suitable technique (Pawlak, 1982, 1991; Pawlak and Slowinski, 1993). Rough set analysis also aims to identify logical statements which can be inferred from a categorical data set and which are the minimally common inferences drawn from such a qualitative data set on various phenomena. In particular, if the aim is to identify one or more factors (characterizing a particular study) which cause the results to differ from those obtained from other rather similar studies (moderator variables), the use of techniques typical of the rough set approach may be particularly relevant.

Finally, if the principal aim is the confrontation and evaluation of the particular methods and the different policies adopted with reference to the same problem in light of the results obtained from the successive individual case studies in order to acquire also some useful guidelines for treating in the future
cases analogous to those studied, an approach based on multicriteria techniques is the most appropriate. Clearly, also the nature of the results and their presentation will depend to a large extent on the technique employed.

The study selection level

The choice of the individual studies to be analysed is difficult and complex. In a certain sense, this phase has the function of an “interface” between the preceding level and the following one, and may to a large extent condition the results of the synthetic meta-analysis study. Although this problem exists for any type of meta-analysis, it is prevailing above all in studies of a socio-economic type, where – as already stated above – it is only possible to speak of “quasi-experiments”, because of the (almost) absence of stable laws governing social phenomena, the use of research methodologies which vary greatly, and the absence of standardization.

From an operational point of view, it is necessary to decide how many and which single studies should be taken into consideration. Methodologically, this requires:

1. defining the group of all individual studies of the given problem which are more or less similar and may be considered eligible;
2. establishing selection criteria among these;
3. deciding on the number of single studies to be analysed;
4. making a selection.

The phases distinguished above cannot be considered as completely separate from one another; for example, if the number of studies considered eligible is reduced, this will have a drastic effect on all subsequent phases. With regard to studies which are potential candidates, a first problem stems from the fact that the individual studies considered eligible are for the most part those already published; furthermore, especially in certain scientific areas, these are exclusively studies in which the results obtained have been positive or confirmative regarding some underlying assumptions. It therefore follows that it is very difficult to take into consideration those scientific studies which may be absolutely correct from a methodological point of view, but which have not reached “positive” results or a clear confirmation (The Economist, 1991; Wachter, 1988). This tendency towards a “continuity” with consolidated results may render meta-analytical research in many fields less objective, since it excludes a priori the possibility of considering results which do not conform to the prevailing line of thought (see Button and Jongma, 1995).

The same problem may also arise for those individual studies of great interest which, irrespective of the quality of the results, are not published, or are published in such a way that they remain inaccessible for meta-analytical review. This is particularly true in the field of applied economics, where, for example, many interesting studies are carried out in the context of private
consultancy or in any case in restricted and confidential circles (the “grey circuit”); at times only some of the principal results of the study are available, without sufficient methodological details (Button and Jongma, 1995).

With regard to the criteria for the selection of the individual studies, particular care must be taken to ensure their similarity; ideally, studies should be chosen which differ in only one fundamental characteristic (by way of controlled experiment), but as this is practically impossible, it is necessary to consider those studies which differ in as few characteristics as possible. An idea of the principal factors characterizing the individual studies is offered in the literature. Besides these intrinsic characteristics, an important comparison regards the different sequences or steps in which the studies were performed (Van den Bergh et al., 1995a). This comparison aims to indicate whether the studies considered are sufficiently similar from a methodological point of view, so that their results may be considered more or less comparable. It is, strictly speaking, impossible to carry out any form of comparative evaluation, classification or synthesis of the results, unless these are held to be sufficiently uniform from a methodological point of view and sufficiently standardized in their presentation.

The principal aim of the similarity analysis of studies is to make a preliminary selection (eliminating those which are too dissimilar) or a preliminary classification (dividing all available and eligible studies and grouping them according to common characteristics). But the real selection of the studies takes place subsequently, taking into account above all the specific problems to be considered, their precise purposes, and the methods and techniques to be used in the synthesis concerned. These are the main criteria to be kept in mind in the final choice of the studies to be analysed. For this purpose, it is easy to understand that the choice whether to employ statistical techniques or, in any case, quantitative analyses – thereby guaranteeing a greater objectivity and a more scientifically solid result – can have a drastic influence on the choice of individual studies. Further methodological problems may therefore arise from the necessity to safeguard the conflicting needs by taking into consideration also studies which, while conforming to the similarity requirements on one hand, may not be suitable for a specific analysis with the chosen technique.

The problem of selection is also closely linked to that of the number of studies to be selected. It is possible to decide on a plausible choice among the existing studies (selection of sub-groups of these studies which best satisfy the pertaining requirements). But it may also be useful (or necessary) to analyse the complete range of individual studies available, or some random samples of these. This choice depends above all on the technique decided on for the study of the synthesis at hand, on the total number of individual studies available and on the intrinsic characteristics of these.

Clearly, it is sometimes relevant to undertake some reclassification of all existing studies on the basis of one or more intrinsic characteristics or other factors that have been defined and evaluated. These classifications may prove to be very useful in the difficult selection phase.
Finally, given the need for uniformity and standardization – or in order to improve the estimation efficiency in the use of quantitative techniques – it may sometimes be necessary to perform further experiments or simulations (Koslowsky and Sagie, 1993) or to carry out new elaborations, calculations or estimations of the data (data manipulation) presented in the individual studies. This need may arise from a finding of technical errors in computing, or from an evaluation of the role of statistical artifacts (sampling errors, unreliability of measurements, etc.) which can raise serious questions about the interpretation of the results (Ones et al., 1994; Witt and Nye, 1992); or from a comparability of research findings utilizing different data collection forms, or in general, from the need to reduce as far as possible the subjectivity inherent in the individual studies (especially in the social sciences) or, finally, from the need to increase the quantity of available data. For this purpose it may be useful to build taxonomies for organizing data, using multidimensional frameworks, etc.

Moreover, in this phase, especially when some quantitative techniques of meta-analysis are used, the problem of the lack of a sufficient homogeneity or the incompleteness of the data available may occur. In other words, an analysis of the individual studies selected may immediately reveal the need to manipulate in some way the data supplied by these studies in order to render them comparable or suitable for elaboration with the chosen methodology. These manipulations, which must in any case be “neutral”, must be performed only if they prove absolutely necessary in order to guarantee the methodological precision and scientific accuracy of the results to be achieved in the synthesis, and in any case must be adequately and explicitly shown.

Furthermore, the problem may occur that some desired data, important for a more detailed and accurate study of the problem to be tackled or absolutely essential for the use of certain quantitative techniques which require this data to be complete, proves to be partially or totally lacking in some of the individual studies (empty cells) (Vanhonacker and Price, 1992). In this case the researcher, bearing in mind the importance of the information contained in this data and the number of individual studies available, must choose between considering only the original data included in the various studies (original design); reducing the data of all the studies, taking into consideration the data remaining after the intersection of these (reduced design); completing the missing data by recalculation, or by estimating it in light of the information available (full design).

The choice between these options is sometimes enforced by a technical need for completeness, or influenced by the advisability of not losing information (considered necessary for the analysis of the phenomenon under consideration). Certainly, when a completion of available data is a result of some hypotheses which are even not substantiated by the more or less explicit information contained in the individual study considered, it is preferable to follow the original design (or, if necessary for technical reasons, to use the reduced design). If, on the other hand, the importance of not losing crucial information is evident, in connection with the suppression of partially incomplete data in some of the
individual studies - as may often occur in problems of comparison, classification, ranking - one may resort to the full design, in any case ensuring the neutrality of the necessary elaborations. In other words, the completion (or, if necessary, the recalculation) of data permits only the use of the chosen techniques, without imposing, however, “options” which are unjustified or not explicitly recognized by the researcher; it is then preferable to obtain synthetic results which are strong and methodologically correct, but perhaps less complete, rather than to supply more detailed and suggestive evidence based on hypotheses of a debatable validity as a result of subjective interpretations.

The meta-analysis level
The principal aim of the work carried out at the meta-analysis level is to conduct a synthesis of the studies selected at the previous level (horizontal transversality). In this phase, the main methodological problem is the one of choosing the meta-analytical technique to be adopted; this problem is also closely connected with the various steps typical of this phase (Van den Bergh et al., 1995a) and, in particular, with the primary objectives of the research programme.

In general, in this important stage, which in a sense constitutes the nucleus of the entire meta-analytical approach, it is necessary to make an accurate analysis of each of the individual studies selected, recognizing all alterations or integrations made, considering the most significant methodological aspects of these studies, identifying the basic results (in relation to the problem faced), and pulling together all results obtained from the analysis of each individual study. These findings must then be compared in order to make an efficacious synthesis (presentation, interpretation and discussion of the results) with reference to the problem faced, to evaluate the consistency and to analyse the robustness of survey conditions; and finally, if needed, to carry out suitable feedback analyses with respect to the previous levels.

To produce this synthesis, it is possible to use a variety of techniques, more or less formalized or standardized (Van den Bergh et al., 1995a). The most important of these are the following: traditional review, content analysis, statistically-based meta-analysis, meta-multicriteria analysis, epistemological analysis (i.e. NUSA P or expert-analysis (Funtowicz and Ravetz, 1987, 1990)), and rough set analysis. The traditional approaches of a literary type suffer from some limitations (Button and Jongma, 1995): outputs in the form of taxonomies without any specific attempt to relate these to the purpose of the review, which constitute merely a description of the problem; a greater subjectivity; no evidence of the degree of conflict of the results analysed; selection of individual studies with similar conclusions (majority rule) without considering the quality of the data and of the techniques used; and serious difficulties involved in mentally handling a large number of different findings.

The use of quantitative techniques, and statistics in particular, can at least partly reduce these disadvantages, ensuring a greater consistency in the results obtained. To use these techniques, however, it is necessary for the results of the
individual studies to be complete and of a considerable homogeneity, characteristics which are not always found in reality, especially not in socio-economic research. In the choice of the technique to be used, it is, therefore, always necessary to bear in mind the principle that it is not possible to generate information which the results of the individual studies are unable to give, while – at the same time – making use of all information that these results, correctly interpreted, are able to offer.

In conclusion, it is possible to say that the selection (and implementation) of the (set of) most suitable meta-analytical techniques cannot be excluded from the context of the whole analysis, but must be made by taking into account all aspects of the problem faced, at each of the levels previously mentioned. Nor is the simultaneous use of different approaches to be dismissed a priori; at times, it may prove extremely useful to carry out the study by using a variety of meta-analytical techniques, even of a completely different nature (Woodside et al., 1993). The failure or success of the use of one of such techniques – and in particular the understanding of the reasons for it – can prove extremely useful, even for a more complete comprehension of the problem faced, in order to highlight sources of divergences in results. For example, sometimes divergent conclusions stem from variations in goals rather than variations in the studies selected to be reviewed.

The implementation level

Also with respect for meta-analysis in environmental economics it is useful to distinguish three different classes of problems on the basis of their goals (see Miser, 1993):

(1) For scientific problems, the goal is to solve the problem, that is, to arrive at intellectual results that are adequate approximations of reality. The function to be performed by the solution is to contribute new results to the field (comparing and ranking studies, correlating aggregated data with other characteristics of each study, identifying moderator variables, etc.).

(2) For technical problems, the function to be performed sets the problem. The task is accomplished – and the problem solved – if the solution enables the function to be performed (summarizing indicators, averaging estimated values and parameters, etc.).

(3) For practical problems, the goal of the task is to serve some actor’s purpose. The problem is solved when a means for serving the purposes has been devised and shown to be effective (evaluating and ranking different methods or alternative policies, aggregating studies by considering complementary results or perspectives, etc.).

At the implementation level possible effects of the study are observed which may go beyond the basic results expected from it.
This phase may assume great importance, especially in some scientific or practical problems where the implementation of a technique, the adoption of a strategy or the suggestion of a practical line of conduct are most desirable.

In general, it can be said that the result to be strived for from any meta-analytical study is a “learning by comparing”, the acquisition of a new experience which is immediately applicable either in the scientific field (for example, design of a new theory, elaboration of the existing explanations, identification of areas in which specific studies or fruitful additional research are needed) or in the application field (for example, developing effective strategies or guidelines for implementation management, indicating types of behaviour consistent with certain objectives, etc.).

Epilogue

In describing the above process, we may use the words of Yu (1990), who claims that an enlargement of the “competence set” is to be obtained, an expansion of the “habitual domain”; these observations lead each individual to a better comprehension of reality and a more fruitful approach to all decision problems, especially those which are new and complex, by forming so-called winning strategies. And this appears to be of particular use in the field of environmental economics, where the policies to be adopted, the decisions to be made are particularly difficult, because of the unpredictability and uncontrollability of processes in the interaction between man and his environment; as a result, it is almost a contradiction to speak of structured decision problems and to require not only the use of suitable highly-sophisticated techniques of decision aid, but also – and above all – of approaches which offer the possibility to understand better the existing relationships between natural phenomena, human behaviour, policy instruments and economic policies. To avoid the trap of methodological complexity and low level information, meta-analysis is a promising approach in environmental policy and impact assessment.

References


