Università degli Studi di Ferrara
DIPARTIMENTO DI ECONOMIA, ISTITUZIONI, TERRITORIO
Corso Ercole I D'Este n.44, 44100 Ferrara

Quaderni del Dipartimento

n.17/2004

October 2004

The Evaluation of Environmental Policy Effectiveness. Methodological Issues and Research Suggestions

by Massimiliano Mazzanti, Maria Gabriella Simeone, Roberto Zoboli

October 2004

The Evaluation of Environmental Policy Effectiveness. Methodological Issues and Research Suggestions

Massimiliano Mazzanti[♣], *Maria Gabriella Simeone*[♦], *Roberto Zoboli*[♠]

Abstract

The present work addresses effectiveness evaluation of environmental policies from a methodological point of view. The aim is to highlight some limitations and opportunities of the approaches currently used in (environmental) policy analysis and to suggest some preferable directions of application based on sets of indicators. The work refers to waste policy mainly as an area of exemplification of the various issues considered.
The work follows a conceptual scheme of policy effectiveness analysis which is based, in essence, on the combination of: (1) Achievement indicators, e.g. percentage of packaging recovered, reflecting policy objectives/targets; (2) Policy response indicators reflecting the (direct and indirect) action by policy-making institutions; (3) Methodological tools by which the role of policy response/action indicators in explaining achievement indicators can be evaluated.
A very general conclusion is that all quantitative techniques have some limitations and policy effectiveness evaluation cannot disregard the information coming form qualitative analysis of policy effectiveness. The latter can also supply very useful information for models building and specification. As a final point, we suggest a research direction aiming at developing detailed system representations as a basis of good and non-ambiguous definition of how policy works and through which channels it arrives to influence the indicators of target/objective. The detailed representation of both the socio-economic system and the policy process is therefore a necessary step in clarifying the causative relationships between, on the one hand, policy Responses and, on the other hand, Determinants and Pressures.

Keywords: Policy evaluation, evaluation techniques, environmental indicators, policy response indicators.

[♣] Dipartimento di Economia Istituzioni Territorio, Università di Ferrara.
[♦] APAT, Roma.
[♠] CERIS-DSE CNR, sezione di Milano.

1. Introduction

The present work addresses effectiveness evaluation of environmental policies from the methodological point of view. The aim is to highlight some limitations and opportunities of the approaches currently used in (environmental) policy analysis and to suggest some preferable directions of application based on sets of indicators. The work is general in nature and it refers to waste policy mainly as an area of exemplification of the various issues considered.

The work follows the conceptual scheme of policy effectiveness analysis proposed in Simeone and Zoboli (2002), which is based, in essence, on the combination of: (1) Achievement indicators, e.g. percentage of packaging recovered, reflecting policy objectives/targets; (2) Policy response indicators reflecting the (direct and indirect) action by policy-making institutions; (3) Methodological tools by which the role of policy response/action indicators in explaining achievement indicators can be evaluated.

The present work specifically addresses the point (3) above, i.e. how to analyse, in a methodologically rigorous way, the *relationships* between environmental achievement indicators and policy response indicators in search for the effects that can be *non-ambiguously* attributed to policy action/response.
The limitations of scope are the following:

- We address issues relating to single specific policies, e.g. directives, and not the evaluation of large-scale policy programmes;
- We do not address single instrument evaluation whereas the whole policy package is of our interest;
- We do not address ex ante policy evaluation whereas we specifically address ex post evaluation;
- We do not address policy impacts in general, for example on employment or innovation, although the latter can be relevant for policy evaluation;
- We are not primarily concerned with the feedback of policy evaluation on policy making.

Two perspectives of our analysis are worth mentioning from the beginning:

- We will refer to approaches based on modelling because models reflect a causative thinking which is intrinsic to structured systems of indicators as DPSIR; however, we will discuss models with a prominent attention to the way they can be used in effectiveness analysis in the case that basic requirements of empirical application of models are lacking, i.e. no modelling tradition, short time series or limited sets of data, little earlier experience in ex post evaluation, etc.
- The assessment of environmental policies has been historically strong on the side of ex ante evaluation; although ex ante evaluation certainly needs further methodological improvements and applied experimentation, we think that the higher "marginal" added value of policy analysis currently relies on the side of ex post evaluation.

We will consider firstly three general approaches to policy evaluation:

- Ex post cost-benefit analysis;
- The experimental approach, based on building a counterfactual;
- The statistical/econometric modelling approach.

Then we will analyse the issues and opportunities associated to three tools often used for policy performance judgements:

- The comparison between actual values of an indicator and a Business-as-usual (BAU) path for the same indicator without policy;
- The analysis of trends of an indicator before and after the application of a policy;

- The analysis of the distance of an indicator from the target level established by policy.

We will conclude by suggesting the pros and cons of the above six approaches/tools and the conditions under which they can be usefully employed for policy effectiveness analysis.

2. Three general methodological approaches

2.1. Ex post cost-benefit analysis

ACB is a widely used tool for policy analysis especially in those institutional frameworks, e.g. the United States and the United Kingdom, where regulation and legislation explicitly provides for the assessment of the impact of legislation itself on the society at large (see Morgenstern, 1997). ACB is well codified by a very extensive economic literature as a tool for ex ante evaluation of specific project and alternative projects comparison. By being the expression of very general evaluation principles, the use of ACB has been increasingly extended to the consideration of very general changes and in particular large scale policy intervention as "projects" suitable to be evaluated with ACB approaches (see Arrow et al., 1996; Farrow and Toman, 1998; Hahn, 1998). Its flexibility allows it to be used in both ex ante and ex post evaluation, either based on very stylised representation of the problem or supported by very detailed econometric models (see, among others, European Commission 1996 and 2000; Risk and Policy Analysis 1998, VV.AA:, 1998).[1]

Model-based evaluation experiences based on ACB approaches are exemplified by the Clean Air Act in the US. In this case, a combination of ex ante and ex post evaluations have been produced. The annual cost of compliance with US environmental laws is estimated at $225 billion in 2000 (EPA, 1997). Freeman (2002) analyses the information on trends in the major indicators of performances of the Clean Air Act (and water laws), over the last three decades, thus adopting an ex post evaluation perspective. Freeman correctly states that in order to assess the effects of the Clean Air Act on emission objectives it is not sufficient to demonstrate

[1] An area in which models have been extensively used for ex ante policy-impact evaluation is climate change. Already at the beginning of the 1990s, there has been a flourishing of models of different formulations, complexity and ambitions addressed to estimate the costs and possibly the benefits associated to different hypothesis of reducing CO2 emissions. Usually they are based on large econometric models employed for carrying out simulations compared to BAU or baseline scenarios (OECD, 1993; Weyant, 1993; Quadrio Curzio and Zoboli 1996; Nordhaus 1992). Models belongs to different categories (e.g. energy models including simplified components of the economic system, or CGEMs including the energy sectors). They are based on cost-effectiveness criteria looking for the minimum cost of achieving a given objective (often defined as the stabilisation of emission at certain levels by a specific date or as emission reduction targets). They simulate ex ante the effects of different control instruments (e.g. carbon taxes or tradable permits) and they employ different measures of cost: the tax requested to reach a specific abatement target the total direct costs, the income compensative variations requested by policies; the GDP losses induced by policies. The costs of policies tend to change substantially according to the different assumption about the recycling of revenue from environmental taxes. The production of economic models of climate change policies had an even greater development after the Kyoto Protocol of 1997, which gave, *inter alia*, a more defined set of targets. The type of model used did not change substantially compared to those summarised above. A stream of model-based evaluations recently arose in connection with the implementation of the Kyoto Protocol in the EU, and in particular about the creation of a European scheme for emission trading (see, among others, Weyant 2000, Bodansky 2002, E3M Lab et al. 2000, European Commission 2000, Biondi, Zoboli, Mazzanti et al. 2002 for a discussion). As an example of ex ante valuation analysis for waste see, for example, Palmer et al. (1996).

downward pollution trends, but it is necessary to compare what emissions and air quality would have been in the absence of the act (in a way similar to the BAU approach, see below).

Freeman then presents an example of ex post evaluation performed on the 1970 Clean Air Act, which is a sound example of ex post CBA, where benefits and costs of reducing emission are monetised (by using both preference based monetary measures and non preference based figures). Benefits are monetised on the assumption that in the absence of the act, total emissions would have remained at the 1970 level (initial benchmark level). Focusing on costs, he considered direct costs associated to economic activities, which generated the most substantial emission reductions. Costs amounted to $43.8, and benefits to $57.3 billions per year.

The U.S. Environmental Protection Agency (EPA, 1997), following the Congress concern about the economic consequences of the Clean Air Act, also undertook a "comprehensive analysis of the impact of this Act on the public health, economy and the environment". EPA then developed a model, which generated estimates of five major pollutants both with the act and what they would have been without it, thus presenting figures on "control" and "no control" scenarios. The no control scenarios show an increasing trend of emissions over three decades, compared to the actual trend of stabilised or reduced emissions. This exercise thus shows that decreases may be attributed to the regulation. The EPA Report, known as "Retrospective Analysis", considers both direct costs linked to economic activities involved in emission reduction, and "other costs", as monitoring, enforcement and R&D by Government. In this case, the benefit-cost ratio is about 28 in 1975, 45 in 1980 and 48 in 1990: the figure thus gives an idea on the dynamic evolution of both figures within the "control scenario". Congress also required EPA to publish an update of those figures every two year, complemented by projections of future figures. This ex post task differs from the regulatory ex ante analysis carried out by EPA in order to assess the need of reviewing existing standards on emissions.

Valuation studies presented by Freeman (2003) show aggregate figures of benefits, calculated by using monetary values (using both individual willingness to pay and avoided damages figures) for mortality and morbidity reduction. They also present a (limited) cost desegregation, in mobile and stationary sources of pollution, and other costs, which nevertheless account only for a limited quota of total costs. The constant rate used to bring figures forward to 2000 is five per cent. [2]

The existing discrepancy of estimates between the two studies (Tab.1) depends mainly on the fact that 75% of EPA estimated benefits stem from reducing premature mortality associated with fine particulates. Then, based on more recent evidence, premature mortality reduced risk is associated to higher values with respect to Freeman first study. In addition, EPA studies assume a greater sensitivity of mortality to particulates matter exposure, and include different assumptions about air pollution levels in the absence of the act. For example, Freeman assumed in his study that, without the Act, the counterfactual scenario was characterised by emissions remaining at the 1970 levels (emission stabilisation hypothesis), a decision leading to likely benefit underestimation.

The above examples of evaluation show the different possible ways of assessing ex post impacts, either by performing comparisons between "control" and "no control" paths, or by estimating costs for different cost sources (cost typology) and in different years, thus depicting a dynamic scenario of policy impacts. The assessment, which is usually carried out by EPA on a full monetary basis, could alternatively be based on analyses attempting to model

[2] The studies are Freeman (1982), Air and water pollution control: a benefit cost assessment, New York, John Wiley, and EPA (1997), The benefits and costs of clean air act: 1970-1990, Washington D.C., Office of Policy Analysis.

relationships/correlations between environmental “performance indicators” and determinants (doses, expenses, policies, etc.).

These studies highlight five critical elements for any ex post evaluation based on the CBA approach:

1. First, all the relevant model assumptions about “with and without” policy scenarios have to be clearly assessed;
2. Secondly, the coefficients of transformation of pollution flows into monetary figures (on the benefit side), if any, should be assessed.
3. Thirdly, along a dynamic scenario, any update of estimates must consider eventual improvements concerning both uncertainty reduction and the understanding of causal links between pollution and health, since both drive changes in benefits and costs figures.
4. Fourthly, at the margin, exogenous determinants like income levels, technological trends and citizen preferences may also have an impact on the estimates of benefits and costs at any given time, to the extent that they affect the value of the environment on the benefit side (mostly increasing it over time), and the cost of reducing pollution flows and stocks (plausibly reducing it over time).
5. Finally, the more the study disaggregates costs and benefits, the more the analysis is able to highlight which specific sources are the most relevant determinants, and what trend (increasing, decreasing, more or less proportionally with respect to key selected variables) is observed for each source over time.

Tab. 1- Benefits and Costs of the Clean Air Act (in billions of 2000$ per year)

Study	Freeman (1982)			Environmental Protection Agency (1997)											
	1978			1975				1980				1990			
	Mobile sources	Stationary sources	*Total*	Mobile sources	Stationary sources	Other*	Total	Mobile sources	Stationary sources	Other	Total	Mobile sources	Stationar y sources	Other	*Total*
Benefits	$0.8	$56.5	*$57.3*				*$468*				*$1.225*				*$1.644*
Costs	$20.1	$23.8	*$43.8*	$7,2	$8,1	$2,7	*$18,0*	$7,7	$16,7	$2,9	*$27,4*	$8,8	$23,5	$2,0	*$34,3*

* Monitoring, enforcement and R&D costs by government.

Source: Freeman (2002)

2.2. The experimental approach and counterfactual analysis

"Experimental methods" are derived from medical or laboratory research and can be used for policy evaluation. The key step of these approaches is the creation of a counterfactual reference for evaluating policy, i.e. a control sample allowing to compare policy application with no policy application. In practice, the "experiment" is based on the definition of a group of individuals affected by a specific policy that are compared with similar individuals not affected by the policy (the control sample). The result of the specific policy is then measured by comparing the performance of the two groups in terms of an "achievement" indicator reflecting policy objectives. The approach involves all the methodological requirements of typical laboratory research but with complication of dealing with individuals and social groups responding to an array of other incentives and decision variables. A good analysis of the counterfactual approach based on creating experimental conditions and control sample is presented in Schmidt (1999). The more detailed methodological presentation is in Heckman and LaLonde (2000).

The use of experimental approaches is spreading in various areas of economic and social policy evaluation, in particular labour policies. An increasing effort on the evaluation of the results of labour policies is being carried out in many countries. The origin of these developments are: (i) at the macro level, the "Luxembourg process" which started in 1997 and imposed Members States to prepare programmes for employment and perform a regular evaluation of the programmes to be reported to the Commission[3] and, (ii) at the micro level, the spreading of so called "active labour policies" that put in place an array of different instruments for improving the performance for labour markets. At the level of application of experimental methods, each kind of "active" labour and employment policy has a more or less developed set of experiences. For example, for professional training policies there have been 18 analyses based on experimental or quasi-experimental approaches in OECD countries during the 1990s (see Favro-Paris 2000 for a survey).

It is interesting to note that, despite the use of these sometimes sophisticated scientific tools, the final evaluation of specific labour policy includes a significant part of *qualitative evaluation and interpretation* of quantitative results (see Antonelli and Nosvelli 2001 for the case of Italy). The combination of quantitative and qualitative analysis is even more significant when evaluation addresses not single policies but very complex policy programmes including various and differentiated objectives. In these cases, scientific tools as the counterfactual exercises are only one of the tools and the final evaluation is much more based on interactive and complex exercises of consensus building among researches, policy makers and stakeholders. Something similar happens with the very complex evaluation process associated to the implementation of the Luxembourg mandate. In this case, also self-evaluation by Member States has a great importance. The other main suggestion emerging from labour policy evaluation experience is that a better understanding of *causal relationships* in the labour policy effectiveness process is more important than to improve standard econometrics applied to labour markets and to experimental exercises.

Recently, the experimental approach based on counterfactual analysis has also been considered in connection with environmental and energy policies, but without significant experiences so far. After stressing the lack of rigorous environmental policy ex post evaluation, Frondel and Schmidt (2001) discuss the relative efficacy of experimental and observational approaches

[3] A similar processes of evaluation of policy programmes have been established in the framework of the Structural Funds of the EU for the Objective 1 regions. The very complex process of implementation of SFs includes ex post evaluation exercises of the income and employment improvements created by the investment programmes.

(statistical and econometric models, see below), suggested as the two pillars of modern evaluation research. They argue that, whenever possible, one should conduct an experimental study, where the randomised control group solves the problem of identifying the counterfactual, once the homogeneity of the two samples is assessed with respect to other determinants.

Nevertheless, environmental policy rarely allows the use of real experiment, mainly when the unit of investigation is "the country" where a policy is adopted (or not adopted). Thus, they point out that "performed appropriately, observational approaches are powerful competitors to experimental studies, not only because experimentation is sometimes not possible" (p. 2). The authors emphasise the need of constructing, in any case, a credible counterfactual situation without the intervention, a scenario that constitutes the ground for a comprehensive scientific analysis. The counterfactual scenario is theoretically a statement on what economic agents would have done in the absence of a policy intervention, and it is usually unobservable. Thus, it has to be replaced by making use of specific identification assumptions or explicit hypothesis[4]. The "non control" sample could be either non-observable, or structurally different, as far as relationships among variables are concerned, from the "control" one. Both issues pose serious problems for evaluation research, but could not be avoided. We will address below how to cope with this issue within different approaches (simulations, regression analysis, BAU analysis, distance to targets, etc.) to ex post evaluation.

2.3. Statistical and econometric modelling approach

2.3.1. A non-technical sketch

Economic and social models are simplified representation of real systems or subsystems useful for testing the existence and strength of causal relationships between variables (or indicators) and to perform simulation or forecasting exercises. This feature gives data availability and statistical/econometric techniques a great role in testing different model specifications. Actually, intensive many-variables testing in search for statistical significance of candidate explanatory variables is often a substitute for a detailed model specification. This too adds to the great variety of models even for the same industry or issue. Theoretical sources of model specification, assumptions, and hypothesis about agents behaviour are very important for large systems (e.g. the economic system) both for including and excluding variables within a very large set of candidate explanatory variables. Even in small models for small subsystem, the possibilities of model formulation are many. Model builders have open and arbitrary choices and they *should* justify those they do. Therefore, a model needs a conceptual framework,

[4] On the possible limitations of such an approach in environmental policy analysis, it is worth quoting from the paper: *"Yet, evaluating policy interventions scientifically on the basis of publicly accessible data with appropriately chosen research methods such as randomised experiments or non experimental difference in differences approaches is unfortunately far from being the standard debate in environmental policy. Current empirical research in environmental economics concentrates on a few selected issues, such as the prediction of emission levels either on the basis of computable general equilibrium models or on the basis of reduced form regressions or the estimation of aggregate production function, with energy as one of the inputs. Thus it seems that further development of empirical strategies in environmental economics is set to follow the course outlined by other research field – rising emphasis of micro data over aggregate time series data, and increasing the use of structural econometric models applied to general purpose individual level data. These changes in emphasis on empirical research occur at a time, though, in which precisely this menu of traditional econometric techniques faces mounting scepticism regarding its credibility. Answers to this sceptic view have included a shifting emphasis away from structuralist to quasi-experimental approaches, the incidental collection of data material appropriate for the research questions, and a careful analysis of measurement issues"* (Frondel and Schmidt, 2001, p. 1).

derived either from formal theory or from literature review (e.g. previous modelling experiences), or from qualitative investigation.

A one-equation model might be written as follows:

$$\text{Environmental indicator}_{1996\text{-}2003ji} = \alpha_i + \beta_{1i}(\text{Vector of socio-economic variables}_{1996\text{-}2003}) + \beta_{2i}(\text{Vector of policy variables}_{1996\text{-}2003}) + \beta_{3i}(\text{Variables representing innovation}_{1996\text{-}2003}) + \varepsilon_i$$

This equation represents a reduced-form model (one endogenous variable, i.e. the environmental indicator, as a function of exogenous variables only) for a longitudinal dataset, where the time span is 1996-2003, and the unit of analysis (i) can be States, Provinces, Municipalities, Firms. The coefficient α_i represent the constant effect, i.e. the part of the indicator not explained by the vectors of determinants[5], the βji are the coefficients of each determinant or explanatory (exogenous) variable, ε_i is the statistical error term in estimation, which is intrinsic to any statistical analysis. Explanatory or determinant or exogenous variables may be grouped in three typologies:

- Socio-economic determinants, e.g. income levels, population density, income distribution, age of the population, share of manufacturing and service firms in the area, etc.; indicators of price incentives on the relevant markets, e.g. exogenous material prices; generally speaking, they can belong to the set of Determinants in a DPSIR framework;
- Variables under the control of policies, in general represented by policy "instruments" (e.g. command and control provisions); the use of quantity based or price incentive mechanism, such as fees, taxes, market permit system, policy-related expenditures); the expected lag between regulatory effort and the changes of the indicator can also be included by dating back the variables; they may be used as "control variables" in order to assess the extent to which specific features of the policy introduced into the system are influencing the dependent variable (the indicator);
- Variables representing innovation; as innovation generally consists of shifts of the behavioural functions instead of movements along the same functions, the innovation variables can be separated form the others.

The model may be able to identify which policy variables and socio-economic variables influence the achievement of the environmental target (Bressers, 1988; Boyd, 2003). The literature on evaluation based on econometric and statistic techniques provides a comprehensive set of guidelines for driving ex post policy evaluation. However, the limits are also quite evident and should be considered case by case. Quoting Frondel and Schmidt (2001, p. 20): "*the fundamental evaluation problem is revealed to be a problem of observability, in technical terms, of identification, not simply of generating larger quantities of unsatisfactory data or devoting more manpower to analysing the data*".

The reduced form model is tractable using statistical and econometric methods even in a cross section of states, regions, firms, individuals. Depending on the nature of data, different regression analysis may be used: (i) binary logit/probit analysis if the dependent variable is codified either as successful (1) or unsuccessful (0); (ii) ordered data if we have categories, plausibly modelled on policy targets (e.g. category 0 for 0-10 percent recycling, category 1 for 10-40, category 3 for 40-70, and so on.); (iii) index variables ranging from 0 to 1 if data availability allows the construction of such "fractional variables" (Jenkins et al. 2003). The simplest way of representing the introduction of a policy in time series models is to use a

[5] The more variables we enter and the more those are significant, the less the costant term is statistically relevant.

dummy variable (value 1 after policy, 0 before)[6]. Policies can be also represented by adding some kind of constraints or by proxies that summarise in just one index the complex details of the policy measure.

The statistical treatment is usually highly flexible with respect to the quality and the nature of data (see Table 2 for a framework of the main steps of quantitative analysis). It is worth noting that, in addition to a sufficient number of observations, the necessary condition to pursue a sound quantitative analysis is the variability of data, across units and/or time. The more variability, either in discrete or continuous terms, is observed, the more significant would probably be the analysis on linkages between dependant and independent variables.

Three elements may undermine the setting up of the model and consequently our representation of reality and the measurement of the relationships: (a) the eventual endogeneity of some variables[7]; (b) the high level of correlation between independent variables[8]; (c) the omission of relevant variables or the inclusion of irrelevant ones[9].

---

[6] Since dummies provide limited information on variability, their significance could sometimes be limited. The setting up of index-type indicators, exploiting all the information on policies, may end up with being a more effective way of introducing policy features into models. Indicators may be shaped as to take into account typical dynamic effects of policies, thus modelling them as increasing from 0 to 1 over a certain time, then eventually decreasing in case the specific policy is changed or expires.

[7] Endogeneity could undermine the explanatory power of the model, both on qualitative and quantitative grounds. If the hypothesis of exogeneity is rejected by data (or by previous results in the literature), the model could become an n-equations system. The point is highly relevant on methodological grounds: both for qualitative (more or less systemic) and quantitative analysis, the assessment of what elements are exogenous (influencing other variables) and endogenous is crucial for the understanding how the system is functioning. The qualitative and systemic approach could inform the quantitative analysis concerning the degree of exogeneity and endogeneity associated to variables. For example, innovation is certainly a variable under suspicion of endogeneity, calling for the use of the other instrumental variables (correlated with environmental innovation but not with other parameters), predictive values for innovation, and other ways of treating an eventual mis-specification. For other variables, the endogeneity can be trickier to assess.

[8] Among the variables selected as explanatory or determinants, some may be highly correlated. High correlation between pairs of variables used as explanatory elements cause mis-specification biases and low quality estimates. As a rule, given a set of exogenous variables showing an index of correlation higher than 0.50, the ones associated to the highest number of significant correlation overall should be dropped in order to preserve the robustness of the model.

[9] Different biases arise when relevant variables are omitted or when irrelevant ones are included: in the former case coefficient are biased, in the second case variances are inflated by using too much information and estimates are less efficient. Thus, the second problem, which we may encounter here in over fitting specifications starting from a conceptual model, is less severe and can be resolved by deleting non-significant variables (using backward selection criteria).

Tab. 2. Steps of quantitative/econometric analysis

Steps	Notes
Defining the unit of analysis (states, industries, firms, households)	The trade off between data availability and sample width should drive the choice on which unit to focus on
Selecting or setting ex novo the index (es) used as dependant variables in the model	Indexes may take the form of discrete variables (binary, ordered, unordered), continuous 0-1 synthetic indicators, or variables taking any real value
Drawing out the set of explanatory independent variables	Divided into different categories, e.g. policy-controlled variables, socio-economic variables, price factors
Defining the vector of policy variables (policy features), by setting up policy indicators, dummy indicators, financial indicators	This is the most crucial step in order to separate out policy-controlled and policy-non controlled variables; both indicators concerning the policy as a whole and indicators of some specific policy instrument may be included
Checking problems of: (1) High correlation among independent variables; (2) endogeneity of explanatory factors; (3) omitted variables	
Choice of the techniques which may be used to empirically analyse the hypotheses on causality direction, which are described by the model	Econometrics is the standard way in economics; other tools may be used as well (i.e. principal components analysis, correlation analysis, etc.)
Compare the results of quantitative analysis with those of qualitative analysis.	Using a model framework single instruments might also be evaluated

2.3.2 Problems in model building for policy analysis

Defining the "right" dependent variable (or achievement indicator)

If the aim of model building is policy evaluation, the choice of the indicator representing the dependent variable, i.e. an achievement indicator, should be made taking into account its proximity to policy action. In other words, some achievement indicators can be relevant in themselves and in terms of general policy objectives but they can be too general for policy analysis because include the effects of many factors outside the control of the policy under examination. In some cases, the preliminary decomposition of the achievement indicator through accounting relationships can help in defining the 'right' achievement indicator that is more directly under control of policy. An example can clarify the point.

Assume that a general objective of policy, e.g. the 6EAP, is to "prevent" waste production or to reduce the production of waste at source (e.g. municipal solid waste from households consumption). At a first sight, the natural candidate achievement indicator could be "waste production" over time (see Kiev background report, May 2003, page 21). However, this is not the right achievement indicator for evaluating prevention policies. Production of waste can be expressed by the following accounting relationship, which is true by definition:

$$PW = C * (PW/C)$$

Where C = consumption in €, and (PW/C) is the intensity of waste production per unit (€) of consumption. By taking the differences of natural logarithm of the variables, the changes over time, the indicator PW can be decomposed into:

$$\Delta \ln PW = \Delta \ln C + \Delta \ln(PW/C)$$

that are now separate and additive. By comparing the change over time of the two components we can identify the share of the changes in PW due to changes of C and those due to changes of PW/C, i.e. waste intensity of consumption. Changes of C are not under the control of waste policy (unless it includes information campaign for inducing families to consume less, etc.). The C variable is actually a very general "Pressure" in the DPSIR framework, not specifically addressed by waste policy "Response". For a given PW/C, changes in C can bring the indicator in many different directions. Therefore, the 'true' prevention policy objective is PW/C, or waste per unit of consumption. In this case, as in other cases, making "prevention" policy evaluation based on the PW indicator can be misleading[10]. Therefore, decomposition of a too general achievement indicator can skip out the part of the observed change of the indicator that surely cannot be attributed to policy, thus leaving only the part of observed achievements that can be related, although in a complex way, to policy action.

The implications are the following:

1. Preliminary decomposition in search for the 'right' achievement indicator is a practice similar to that used in "decoupling analysis" and environmental Kuznets curves (see, among many others, Borghesi and Vercelli, 2003), where scale and efficiency indicators are separated and the latter are the most significant for *specific* policies;
2. Decomposition of the observed changes of an achievement indicator that indicate the components independent from policy and those under the influence of policy, can be a very simple way for a preliminary evaluation of policy. In our example, given a change in the PW, if a substantial share of the total observed change is due to the policy-controlled variable PW/C then it would be a general suggestion of policy effectiveness.
3. In our example, PW is an important indicator in itself, and it is the best indicator for measuring the pressure on the treatment capacity and natural resources at large, but it is nevertheless not a good indicator for effectiveness evaluation of policies aiming at "prevention of waste production". In this case a relatively better indicator is the ratio PW/C. As a consequence, if the achievement indicator addresses is PW, we are not looking at prevention policies *only*, and instead a set of policies, not necessarily limited at waste policy, would be actually under scrutiny.

[10] Suppose that the actual value of PW is decreasing over time compared to the base year, and/or the past trend, and/or a BAU scenario: it sounds like a policy success. However, decomposition of a decreasing ΔPW over time suggest either: (1) ΔC over time is negative and $\Delta PW/C$ is zero: all the policy success depends on a factor not controlled by policy, the exogenous determinant; but even in this case, we cannot affirm whether the policy was unsuccessful because the stability of PW/C can derive from, for example, a collapse of packaging material prices that brought to increasing packaging per unit of production/consumption, and policy might have reduced this adverse market factor successfully; therefore, a further analysis of the variable/indicator under the scope of policy, but not entirely governed by it, is necessary; (2) ΔC over time is positive but PW/C is strongly decreasing and this apparently a success of policy; however, policy cannot be said to be successful unless other factors are examined: for example, a strong jump in paperboard price might have induced companies to reduce the intensity of packaging, and policy might have done nothing good for achieving the "prevention"; here again analysis is necessary looking at non-policy variable acting on variables within the scope of policy action.

Relationships between policies and non-policy explanatory variables

Following the sketch of modelling approach, in a reduced-form model explanatory variables should reflect the effects deriving from:

1. Exogenous policy variables;
2. Other exogenous variables not influenced by policy variables.

However, in various policies, there is the possibility that some of the non-policy explanatory variables are actually not independent from the policy variables. For example, in the case of PW/C indicator above, it can be explained by (in brackets the sign of expected effect):

1. Per capita GDP (high income consumers produce more waste per unit of consumption, unless there are decoupling and/or "Kuznets-curve" effects) (+/-);
2. Product innovation in mass consumption good (product differentiation by advertising packaging, added service through packaging) (+);
3. The share of consumption outside home (+/-);
4. The market cost of packaging for manufacturing and distribution system (-);
5. The policy-induced cost of packaging through a recycling fees (-);
6. Taxes on households for municipal solid waste (-)
7. The availability of collection facilities (+/-)
8. Investments done on consumers' awareness, e.g. information campaigns (-).

In this (reduced form) equation for PW/C, only variables from (5) to (8) are under the partial or full control of waste policies and their instruments, i.e. policy action/response. However, some *indirect effects* of policy can be missing in the reduced form. For example, in packaging policy, the packaging cost to producers and distributors is *directly* influenced, i.e. *increased*, by a recycling fee imposed on packaging, but the same policy instrument can have the effect of increasing the supply of packaging materials to be recycled (as is the case with the schemes as DSD in Germany or CONAI in Italy), thus *decreasing*, ceteris paribus, the market price of packaging materials, another variable in the list[11]. Thus the need of examining correlations ex ante. The indirect effects (having different sign compared to the direct one) cannot be seen outside a complete model of the market for packaging (or for waste in general) at least including equations for "supply" and "price" of materials from packaging waste. In the supply equation, the policy instrument "recycling fee" can be directly included as explanatory variable then entering the reduced form equation for waste packaging price. The latter can enter an equation for packaging material price together with international virgin materials price.

Furthermore, the variables under the control of policy included in the list above may influence the indicator as a side effect of other policy objectives not represented by the indicator equation and their role can be ambiguous. For example, the creation of collection facilities for MSW is an essential objective and sometime an instrument (through investment financing) of "sustainable waste management"; but the availability of good collection facilities can reduce the cost of producing waste for households, thus reducing the incentive to reduce waste intensity of consumption[12].

---

[11] While the recycling fee may be exogenous, market price for packaging materials is not exogenous because it is partially determined by the recycling fee itself through impulse on supply of waste materials to be recycled; a possibly exogenous variable in the "reduced form" will be international price of virgin materials (instead of national cost of packaging materials); similar issues arises with the so called "rebound effects" of environmental efficiency.

[12] The latter effect can be even more true in the case there are policy schemes giving money to households for their effort in delivering waste at separate collection facilities for e.g. improving recycling, which is another objective of waste policy; in essence, the ambiguity of the sign of the effect is greater the greater the simplification represented by the reduced form equation.

The main implication is that a good representation of policy *direct and indirect effects* is preliminarily required. Before arriving to a simple reduced form suitable for econometric estimation in search for partial effects of policy variables it would be necessary to build up detailed models of the system.

Problems in defining policy variables

The issue of selecting or building indicators (variables) representing the action of policies is more difficult that the selection of achievement indicators. The choices made for the case of packaging policy evaluation are presented in Mazzanti, Simeone, Zoboli (2003). Here we give three examples for these difficulties: (1) the lagged effects of policies; (2) cost and expenses as policy response indicators; (3) the possible role of innovation as a policy variable.

(1) *Lagged effects and effectiveness cycle.* Sometimes, the disappointing results of model-based policy evaluation are explained by the fact that implementation of policies requires time and the formal representation of lags is difficult. There can be a systematic uncertainty on whether the (ir)relevance of policy variables derives from their (in)effectiveness or from a difficulty in measuring the lags of impacts. In general, there is a "policy cycle" - from first proposals to adoption, implementation, and reactions by involved actors, etc. - that has a great importance for effectiveness and its evaluation. Model based analysis of policy cycles can be difficult unless there are favourable conditions, in particular a long-enough time of application, enough available data, known factors allowing to represent the main the way policy affected the system. The latter, i.e. representation of policy history, should be a preliminary step of policy effectiveness analysis. Obviously, if the latter takes place at the beginning of policy implementation process, it can address only few elements, e.g. the role of expectations that the discussion of a policy measure created for some, e.g. industrial, actors (this is the case with many directives at their initial stages). The attempt to define a typical policy cycle composed of different phases (for directives: proposal, discussion, adoption/transposition, etc.) can be undertaken even arriving at estimating average lag times for each phase across different regulations of the recent past or average lag in implementation and of the same policy in different countries. Such an analysis can provide a representation of cycles similar to an "investment cycle", with a first phase marked by high investments and very low "returns" (low achievement indicators) and subsequent phases marked by higher "profits" (high achievement indicators) even with stable or declining investments[13].

(2) *Costs vs expenses.* A candidate variable for representing policy action can be direct and induced cost for policy implementation. Expenses seem to be a proxy for "costs". However, expenses and costs show different perspectives: expenses are closer to private and public investments, thus representing a close and instrumental consequence of policy action. Instead, costs are referring to all figures of direct, indirect and shadow costs (opportunity costs) associated to policy implementation and compliance with the policy, by both private agents and

---

[13] The existence of policy cycles, not easily represented by formal modelling, raises a further question about the partial exogeneity or endogeneity of policy variables. Not only policy introduction is a process evolving in time along a non-deterministic cycle, but there are also various mechanisms pushing to consider it as partially endogenous. The great role of the discussion between stakeholders and policy makers in making regulations is well known, as is the issues of "regulation capture", asymmetric information between industry and policy making, and other issues. The latter creates dynamic interactions between the system and policy in some phase of the policy cycle, not allowing a definition of policy as truly exogenous. This has great importance for effectiveness evaluation. Some policy provisions or full policies might be born with objectives and targets that are achievable with a low effort. In other, words, the greater the endogeneity of the policy formulation, the greater the ex post effectiveness we can expect to observe.

eventually by society as large (if social market and non-market costs are also accounted for). Therefore, costs can also be accounted for as a part for the "achievements" of the policy (although with a possible negative sign) that parallel other achievements on the environmental side. Expenses may also be used with some caution in cost-effectiveness analysis of regulation/policy, provided the different units show the same level of performance indicator. Otherwise, the assessment is just possible on the basis of an examination of the effort specifically devoted to the environmental program under scrutiny. No efficiency consideration is instead plausible. What costs to include may represent a final controversial point, which is to be investigated case by case[14]. Financial costs, current and capital expenses, indirect costs, external costs, opportunity costs are all possible candidates to enter ex post evaluations.

(3) *Innovation as a policy variable.* An issue that can bring serious mis-specification problems in modelling is the representation of innovation either induced by policies or taking place as the result of normal business. Innovation can be an intended instrument or intermediate policy objective as in the case of IPPC directive based on BATNEC. In general, however, it is an unintended impact that, however, can be critical for policy success and, at the same time, unpredictable. The analysis of innovation induced by environmental (or dynamic-incentive impacts) has been an expanding field of research during the last few years[15]. Depending of the technological profile of the addressed sectors, dynamic-incentives assume a critical role in policy effectiveness analysis. As suggested by an increasing number of works, technological and organisational innovation is actually the main 'response' to environmental policies addressing emissions and waste from industrial activities whereas the level of activity is little affected. Among the various problems in representing innovation in policy analysis, a crucial one is the transmission of policy incentives in presence of innovative reactions by economic actors. Innovative choices, both induced by policy and autonomous, can follow non-smooth patterns based on discrete choice (yes/no) according to critical cost/benefit benchmarks or, even more elusively, can be governed by industrial strategies. In this case, expectations about the effects of, for example, economic instruments my be frustrated. The use these instruments corresponds to the assumptions that a cost/price incentive allocated to a specific stage/actor will be "transmitted" through markets to other stages/actors and, at the end, the "right" or desired innovation path will prevail. However, economic instruments modify the cost-benefit balance that actors associate to innovation choices and the result can be unpredictable depending on the specific industry/activity, its market relationships with other industries, the technological and organisational capabilities of these industries. Therefore, the analysis of the innovation processes and possibilities involved by policy should come first relative to policy effectiveness analysis in search for the right representation of exogenous and policy-induced innovation contributing to policy effectiveness[16].

3. Three specific tools for effectiveness analysis

We analyse below three specific tools often used in producing judgements about (environmental) policy effectiveness: (1) comparison between actual values of an indicator and

---

[14] As stated by Freeman (2002, p.133): "*Another issue involves the omission of indirect or general equilibrium effects in the estimate of costs. The EPA's cost estimate is the sum of annual direct expenditures on operation and maintenance and the amortised capital investments in pollution control equipment. Not included are the indirect costs that arise through general equilibrium effects in labour and capital markets that are already distorted by income and other taxes*".

[15] See, in particular, Jaffe et al., 2003. For analyses on Europe see Hemmelskamp, Rennings and Leone (2000), Hemmelskamp and Leone (1998), Klemmer (1999), Kemp (1997), Mazzanti and Zoboli (2005).

[16] For example, by estimating an equation for innovation, then using the result in policy evaluation equation.

the values it would have had without policies, or Business as usual; (2) the analysis of changing trends of an indicator after a policy has been introduced; (3) the measure of the distance from policy targets for an indicator as a suggestion of effectiveness. We try to highlight the limits and the possible use if the three tools also in relation to modelling and the methodologies discussed above.

3.1. Comparison with BAU

An often proposed framework for evaluating the impact of a policy is the comparison between the actual evolution of an achievement indicator reflecting the policy objective (e.g. production of waste), and a possible alternative path that would have occurred *without* policy (or with a different policy, see below), which is usually defined as Business-as-Usual (BAU) scenario or path (see EEA, 1997, 2001, 2002; ETC-ACC 2003).

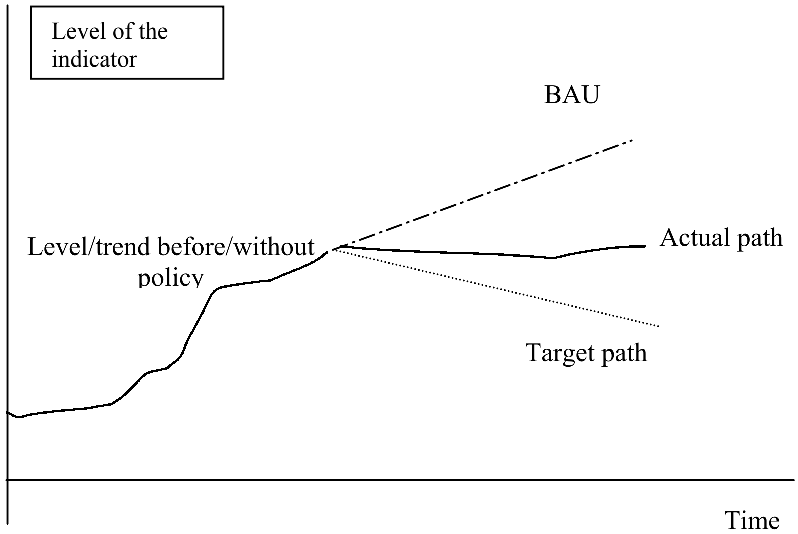
In general, the comparison with BAU (usually done for time series) can be considered as logically equivalent to the "control sample" used in the 'experimental approach' (in a cross section or panel setting)[17]. It is very important to note that, when it is used in ex ante policy analysis (as climate policy models), BAU is a simulated path of future non-observed values. In ex post evaluation, instead, BAU is calculated for a period on which *actual observed data are available* for the indicator and it is a simulation of values the indicator *would have taken* without policy application, a path which is obviously non observable.

To illustrate the usual BAU framework, we assume that the history of the problem addressed by policy can be represented and measured by an indicator (e.g. waste production). It can be compared with respect to: (a) the historical level of the indicator before policy was introduced; (b) the simulated path without policy, or BAU; (c) the quantitative target of policy, if any, in terms of the indicator (see Figure 1, where BAU is from a linear projection of past trend before policy).

The comparison of actual values with the past trend before policy, i.e. the (a) element, will be considered in Par. 3.2. Not all policies have the element (c), i.e. a quantitative target for a certain date. If they have it, the path-to-target is generally represented by a straight-line from the (policy defined) initial state to the target point. This case will be considered in Par. 3.3. Here we consider only the comparison between actual values of the indicator and BAU path.

[17] Actually, a "control-sample BAU" could be defined with large samples of countries or firms or individuals including both those affected by policy as well as those not affected.

Fig. 1. The BAU framework



For ex post policy evaluation of a newly introduced policy, BAU should reflect the effect of *all* variables influencing the indicator *after* policy introduction *excluding only* the effect of policy variables, so that the difference between actual path and BAU path represents *only* the effect of the policy. Therefore, the main shortcomings may arise with the ways the non-observable BAU is possibly estimated. There are basically two ways a BAU path is calculated in empirical studies:

1. by projecting past trend using statistical techniques, which can be either simple, e.g. the regression of the past series on t and then projection using increasing values of t, or more sophisticated techniques for time series analysis (e.g. ARMA-type models);
2. if there is an econometric model for the indicator, by estimating the model *excluding* variables representing policy action.

In the case (1), BAU does not exactly respond to the required properties of a counterfactual in which the only change is the introduction of policy. BAU path deriving from a projection of past trend is the result of a set of effects and its comparison with the actual path of the indicator can be misleading for the evaluation of policy-effect[18].

In case (2), model ability to fit with actual values of the indicator before policy introduction has a critical importance for the results. With a model, three paths actually enter the comparison: (i) recorded actual values; (ii) actual values *before* policy that are *estimated* with the model; (iii) e*stimated* values *after* policy introduction but excluding policy variables, i.e. the BAU path[19].

As a first requirement, the model should fit well with actual values *before* policy, otherwise comparison of actual values (*after* policy) with a BAU (*after* policy *without* policy variables) calculated with that model will be biased.

[18] For example, during the application of a policy other factors might have coincidentally entered into the fore that could have changed the past trend as reflected in BAU even in absence of policy; in a actual-to-BAU comparison that effect would be attributed to policy.

[19] In econometrics, BAU is technically equivalent to a so called "ex post forecast" estimated by using parameters estimated over a certain period (before policy in our case) and the values of exogenous (non-policy) variables in a subsequent period (after policy introduction in our case).

As a second requirement, exogenous non-policy variables after policy introduction, used in estimating the BAU path (with the coefficients estimated before policy introduction), should remain independent from policy introduction. If policy introduction should change the relationships between explanatory variables or the coefficients of exogenous variables with respect to the before policy estimate, the BAU calculated with exogenous variables after policy would incorporate some effects of policy itself, and the comparison with the actual values after policy would be misleading.

The consequence is the need of good model specification in terms of exogeneity of variables in the estimation before policy. Limitations arising from bad model specification in terms of exogeneity of variables, can be exemplified by looking at variables such as technological innovation (see also above). If the model includes innovation variables it is critical that they are independent from policy. If they are not, i.e. they are fully or partially policy-induced, their actual values after policy, which are used in estimating the BAU, will be influenced by those policy variables that are excluded in BAU estimation. In this case, the estimated BAU path still incorporates some (indirect) effects of the policy.

More in general, policy effects might be so pervasive that relationships among exogenous variables and estimated coefficients for the period before policy could be no more valid after policy introduction, but they will influence those actual values of the non-policy variables used in BAU estimate. In other words, when comparing BAU and actual values, we might actually compare not the same model for two periods but two completely different models, the model after policy reflecting a pervasive policy-induced change of correlation among exogenous variables and different behavioural parameter compared with that before policy.

Take for example a country where a policy such as the DSD system has been implemented for packaging. The ‘model’ for the packaging waste system before and after DSD can differ by:

1. a different right of households to dispose off their packaging waste;
2. a different legal responsibility of packaging and industrial producers;
3. an economic instrument like the ‘green dot’;
4. a possibly different operational setting of municipalities in managing MSW;
5. a new set of commercial agreements between recyclers and the bodies implementing packaging policy.

All these differences correspond to an array of policy-induced changes that are not reflected only in smooth changes of quantities and prices of the packaging-related ‘markets’, but can bring to changes of the *whole* behavioural functions of households, packaging producers, municipalities, recyclers, etc as represented by the parameters or elasticities. For example, in the Italian case, a BAU path for an indicator like ‘landfill of packaging waste’ after CONAI introduction in 1998 should be calculated as if the behavioural parameters of all actors prevailing with the old system based on *Consorzi Obbligatori* and the exogeneity of variables would remain the same after the introduction of CONAI, then assuming that the latter is a ‘marginal’ change not affecting the structural stability of the parameters estimated before 1998 and relationships between variables. This is obviously not realistic and a BAU estimated with a model in such setting will reflect exogenous variables that might be no more exogenous and behavioural parameters (e.g. elasticities of supply to packaging prices that are completely changed due to CONAI system.

A proper use of BAU path comparison requires a certain degree of structural stability of the system representation, which is not the case with some ‘whole-system’ policies. A more appropriate use of BAU would be, therefore, for simulating the effect of specific policy measures that do not alter the system structure or for simulating different levels or details of the policy variables. To test stability and its degree, we can suggest to estimate parameters with and without the policy package, i.e. using non-policy variables before and after policy, and to compare coefficients/parameters resulting from the policy-with and the policy-without models[20].

Another issue arises when the comparison of the indicator path under policy action is not with a BAU without policy, but with a BAU that incorporates *another policy* existing before, which is subject to revision of objectives and targets (i.e. comparison between, say, Policy 1 and Policy 2). This can be the case of packaging policy having old and new targets. In this case, BAU should be a 'constant policy' simulation of what would be happened by going on with Policy 1. Provided that the above consideration remain valid also in this case, may be that the change from Policy 1 and Policy 2 is not so radical as is the passage from no-policy to with-policy situation. Therefore the new policy might not represent a too radical change and the structural stability of the system could be greater.

However, additional issues may arise in connection with *lagged affects of policy*. If there are lagged effects of Policy 1, and the model is not able to take into account them, BAU will not generally include these effects and the estimated BAU path can be distorted. On the other hand, also actual development will include the lagged effects of Policy 1 during the application of Policy 2 (e.g. the results of long term investments in plants and equipment), and the path representing Policy 2 is biased by these effects (possibly not included in the BAU with which it is compared). Furthermore, lagged effects of Policy 1 can interact (not additively) with the effects of Policy 2, which cannot be reflected in BAU. In essence, in the presence of lagged effects, the risk is to attribute lagged effects of Policy 1 to the action of Policy 2.

3.2. Trends before and after policy and break points

The problems in estimating a BAU path with models suggest not to rule out other non-BAU forms of analysis as the comparison of the indicator trend before and after policy. The simple comparison of the actual level of an achievement indicator with the initial state or before-policy trend is widely used in indicator analysis and reporting although it is apparently meaningless for evaluation in terms of policy effects alone. However, there are conditions under which the comparison of present levels with those prevailing before policy is meaningful[21].

A comparison with past (before-policy) trends can be useful when there are very long time series allowing to define (though statistical techniques) reliable trends before and after policy, and then the structural breaks. It is clearly a "black box" analysis not giving a direct explanation of the exact role of policy but it can be useful if the achievement indicator is directly and fully under the control of policy, so that the attribution of a role to policy in breaking trend is reliable. In other words, it answers the question: did policy introduction or change break the trend of the indicator? If yes, was the change significant or not? This is probably the case with command and control approaches. For example the prohibition of disposing off untreated waste by landfill within a certain year must give rise to the decrease of the achievement indicator "non-treated waste in landfill". A long time series of the latter indicator can be suitable for studying the effectiveness of this policy by searching (by statistical techniques) structural breaks in the trend before and after policy. We expect that a significant lag will take place before full application and target achievement and we can search for factors (e.g. lack of pre-treatment facilities) explaining this lag of effectiveness. Nevertheless, comparison with trends before policy provision can be meaningful.

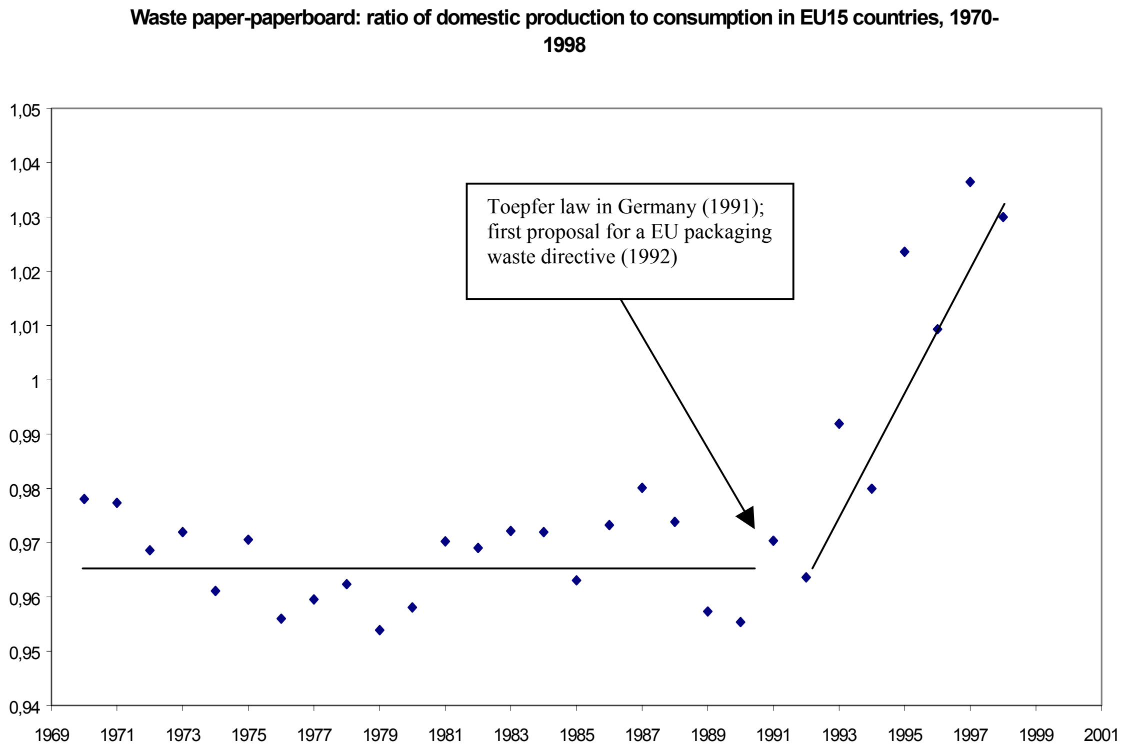
[20] In order to allow a sensitivity analysis of the difference between the two possible BAU estimated trends.

[21] In particular if the alternative is a BAU calculated as a projection of past trend: the comparison between the present trend and either a past trend before policy or a BAU as a projection of past trend are not too different.

The presence of structural breaks associated to "policy shocks" may be assessed by specific tests (Chow test, Fisher-type tests, Hausman Test, Likelihood ratio type test), provided data are available over a sufficiently long time[22]. The intuition behind these tests is that the hypothesis that model assumptions apply to all the observations may be incorrect. A structural change may emerge when dividing the sample in two or more parts. If the test confirms the structural change, the two samples should be analysed as separate entities, and the policy introduction would end up in being effective in changing the indicator trend. Coefficients linking indicators and determinants are then different. It is clear that qualitative analysis should inform about the possible points of structural change within the dataset. For our framework, the assessment of structural changes could in itself be presented as a proof of significant policy impact; then, it remains to verify the impact after the policy implementation. It remains that the policy change to test is defined by researcher ex ante on qualitative grounds.

Trends-and-breaks analysis can also be useful as a first exploration of policy-effectiveness hypothesis with highly stylised indicators. An example is presented in the Figure 2 below where the ratio of waste paper-paperboard production (i.e. collection) to consumption for the (present) EU15 countries is depicted from 1970 to 1998.

Figure 2



Source: calculations on FAO data.

The ratio can be also interpreted as the degree of self-sufficiency (lower than 1 = net importer; higher than 1 = net exporter). It suggests that, after a very flat trend for almost 20 years, at the beginning of 1990s the collection of waste paper-paperboard in EU countries rapidly began to

[22] See for example Greene (2000, p.287), who presents a case about gasoline prices and per capita consumption over the period 1960-1995.

grow more than consumption. There is a clear break compared to the past (which is not tested here with statistical techniques). The early 1990s have actually been the years in which packaging waste policies began to take place in various countries, e.g. the Toepfer law in Germany in 1991, and in 1992 the Commission made the first proposal for a packaging waste directive. The suddenly changing trend of the above indicator might then be considered as a general suggestion of policy effectiveness in the packaging waste sector (although the indicator includes both paper packaging and waste paper from MSW). Nevertheless, it must be noted that the denominator of the indicator (consumption) is the input of waste paper-paperboard in recycling processes and the indicator suggests that the domestic capacity of recycling did not grow in line with a rapidly increasing collection, thus giving rise to increasing international trade flows of waste paper. The latter developed firstly among EU countries and the EU as a whole became a net exporter of waste paper to many destinations outside EU. We should conclude that the possible effectiveness of packaging and MSW policies has been for collection of waste paper but not for recycling to the same extent. Therefore, evaluation requires a deeper analysis and the study of changing structural trends can be only the starting point.

3.3. Distance from targets

In the case policy targets for the achievement indicator are established, the analysis of distance form target at a certain point of time represents another extensively used way for formulating judgements about policy effectiveness, for example in climate policy[23]. It can be performed both within or outside a BAU framework (see Figure 1) and both during time and across countries.

During time

In a non-BAU framework, distance to target analysis is apparently trivial in that it consists in defining how actual developments of the indicator/variable subject to target are performing compared to the (linear) path towards target. Its use in monitoring can be that of an alerting measurement device (see the EEA analysis on distance from the Kyoto targets in terms of temperature in a thermometer) and suggest similarities between policy and dynamic (automatic) control processes (see Par. 4).

In a non-BAU framework, the evaluation through distance from targets comes from two possible joint comparisons: (a) actual level at a certain point of time compared with the initial level (starting year of the target or starting date of policy introduction): (b) the position of actual level compared to target(s) trajectories. The first comparison (with initial level or the trend "before-policy") can follow the same lines discussed above about the search for breaks and changing trends. The interpretation of position with respect to target path can actually implies two different questions: (a) how much are we distant from the target and how much is yet to do in the remaining time?; (b) are we on a convergent or not convergent path with respect to target line? The two questions are different for policy evaluation. The first one asks what remains to do and might not imply a discussion on the effectiveness of what we have done, although it may stimulate an evaluation on the effectiveness of the change compared to the initial state. The second one implies a question about the effectiveness of what we have done (are doing) compared to what we should have done for convergence. The second question is less simple because, if policy effects need time to have effects, then how can we be sure that we are on a convergence or non-convergence path? And how can we confidently re-address our policy to be more effective?

[23] Target definition and adoption, especially if they are binding and close in time, is a much debated issue of environmental policy making (interaction between economic agents and policy makers, bargaining over percentage points of reduction and/or years of application, and so on). The above arguments about a certain degree of policy endogeneity therefore apply.

In principle, having a model for the system which includes estimated lags of explanatory variables, we could be able to estimate a convergence path (different from the simple straight line of target path) depicting *where we should be* at each point in time for being converging, then estimating our ineffectiveness degree and the role of the different variables (including policy variables) in explaining ineffectiveness[24]. The possible complexity of such an exercise raises all the questions about models' use already examined in general and with respect to BAU. All in all, a straight line target path from initial point towards targets is probably the best approximation of the desirable path for comparison in convergence analysis.

Across countries

A different interesting possibility is to consider a distance to target analysis across countries. By exploiting some suggestions from convergence and catching-up analysis developed by research on international economic growth during the last twenty years, the distance from target analysis could be useful for cross-country comparison of policy performance. Considering policies with the same targets across European countries, the typical configuration during time is that some countries are more rapidly approaching targets than other countries (not surprisingly, target differentiation is a significant issue of debate in some environmental policies). This might allow to address not the absolute but the relative ability of countries in achieving the target. This can be a form of benchmarking. For example, assuming that the beginning of a policy is the same in all countries, the distance to target for the achievement indicator for the best performing country (the more close to the target at a specific date) can be the reference for measuring the distance of each other countries at the same point in time (a vertical distance on the time axis). If there is a lag of policy start in some countries, this can be accounted for in measuring distances. Over time, we can obtain a good set of information on the *relative* speed of achievement. Moreover, some suggestions about the feasibility of single common target Vs the desirability of differentiated targets may emerge.

Obviously this kind of exercise presents some limitations and risks in terms of policy effectiveness evaluation. The immediate interpretation would be the relationship between speed of convergence to benchmark and the form the policy has taken in the different countries, thus making inference on the superiority of some “national” policy approaches. This would be risky because an array of other condition may favour some countries in achieving the targets before[25]. Therefore, following the suggestions of economic growth convergence analysis, a set of variables can be used to test econometrically the explanatory variables of different country performances and to identify the role of non-policy (local) conditions as well the role of different policy approaches.

Although it is an evaluation subject to some limitations, it can suggest the relative effectiveness of a policy package as a whole (estimated as a kind of residual after defining the role of non-

[24] In practice, for policy evaluation purposes, it could take the form of an ex post simulation exercise in which the target is used to estimate the “desired” level of all model variables up to now, and then the level of both policy and non-policy variables required to have those values. As non-policy variables can take whatever value, the result is a normative suggestion on which levels of policy would have been needed to have the “desired” levels. The latter can be compared with those level the policy variables have actually taken and the distance can be a suggestion of effectiveness/ineffectiveness, as well as required policy adjustments. The main problem is associated with the residual time for achieving a target, the period from now to the final implementation stage, because we do not know the non-policy variables affecting the future path and to what extent we have to adjust policy variables for: (a) recovering the possible ineffectiveness of the past; (b) increasing effectiveness in the remaining time to targets.

[25] Including the fact that some countries might have been better positioned in influencing target definition in the European policy making, i.e. endogeneity of policy.

policy variables) instead of specific policy instruments. This leaves the problem of interpreting *why* some policy packages are successful and others are not. An interpretation problem, however, is that inter-country differences in policy approaches can be so important compared with non-policy variables that almost whole difference in performance might be attributed to them. But the approaches differ largely because of the institutional, social and economic situations of the countries are so different that some approaches would not be feasible outside that specific country, and the final evaluation of superiority of some approach is useless for policy implementation. In other words, policies may be partially or totally non-transferable.

For the feasibility of such a cross-country (or cross-industry) analysis, given the dynamic setting which characterises the achievement of environmental targets, a longitudinal dataset would be the best information framework for analysing the impact of set of explanatory variables[26]. The longitudinal nature of data would partially solve the problems when dealing with cross section data, basically the theoretical definition of causality direction. The main problem concerning panel data is nevertheless the necessity of setting up a full dataset, which must necessarily rely on a sufficient number of periods (year), and a sufficient number of units. Obviously, in country-level analysis, the available number of years associated to reliable data should be large in order to perform a consistent econometric analysis. Using firms or industries, the unit constraint is less restrictive, since it is easier to collect information even on a large number of agents.

For example, the implementation of the 94/62/EC Directive on Packaging will witness at least a 12 years time (1994-2006) of policy application, implementation, target re-definition and final evaluation. Information on twelve years and fifteen or more countries would lead to a sufficiently relevant sample for statistical analysis. Now, the maximum amount of data potentially collectable derives from waste generation and treatment reports by 15 EU countries over the period 2001-1997. This would mean a 15x5 matrix dataset, not a large enough dataset indeed, but sufficient for a pilot quantitative study. In 2006, a 15x10 matrix could allow an analysis associated to a sufficiently robust statistical significance. Other EU directives on environmental issues, whose implementation period is longer, would provide even larger dataset at the final stage of evaluation. In cases where information concerns (and is available) for many units and many periods, statistical analysis may be also used for intermediate evaluation.

4. A suggested direction of research

All the approaches and tools examined above highlight that policy impact or effectiveness analysis needs a good definition of the *causal links* among variables or indicators. When theory about behavioural functions of agents is still weak and datasets are still poor, which is still the case for various environmental policies, we should use more detailed descriptions of relationships among actors and indicators to discover the *process* by which policy is effective or not. In other words, system representation in terms of detailed causal links between variables can be a good tool for developing policy analysis when statistical techniques are only partially useful due to limited information and there is not the possibility of reliable econometric estimates over large samples of data.

The opportunity of looking at detailed representations of relationships among indicators also arise from the fact that decisions by actors in environmental sectors are often dichotomous, i.e. decisions to do or not to do according to critical values, e.g. opportunity costs of landfilling, or

[26] Although even a cross section environment may capture dynamic linkages if lagged information is used to explain target achievement at current time. This also helps dealing with the intrinsic problem of reciprocal causative effects.

according available capacity for recycling in not-too-distant locations, or according to specific legal provisions not allowing to do something (i.e. some variables can be bound by non-linear constraints and some others can be subject to non-divisibility). In these cases, the representation of actors behaviour according to smooth continuous functions of a set of variables, which is typical of traditional modelling, can be inconsistent with a reality in which policy itself created discontinuities and multiple-constraints decision processes.

Explicit representations of choice processes through mixed mass-balance, engineering, microeconomic and institutional relationships, even using dynamic specifications and calibration parameters derived from case studies, can allow to depict and track the causation chain of policy action. Conceptual references for treating detailed system representations of such kind can be control and feedback theory, dynamic adaptive systems (from Lange 1977, to Day, forthcoming) as well as industrial dynamic models. In these approaches, policy can thought, in general, as a control device pushing the system towards a desired state, as represented by objectives and targets, through its influence on technical and behavioural relationships. Indicators are the variables on which the dynamics of the system is monitored and measured. The similarities with dynamic control models (e.g. temperature control devices) can be straightforward, as can be seen from various dynamic-simulation packages.

The waste sector can be a good area for developing such an approach. The waste-related system (waste production by households and industries, final disposal, recycling/recovery, etc.) represents a complex set of actors, technologies, and markets linked together by material flows, economic flows and institutional interdependencies, *many of which established by waste policy itself*. The increasing information on such a system in all countries supplies the opportunity to build detailed system representations that include the relevant links and the definition of causative influences. The latter can include technical, engineering, and ecological relationships together with economic and market relationships, largely shaped by the existing regulation, legislation, and public administration practice. The open nature of the system, for example through recycling which involves an array of industries producing/consuming materials, do not prevent from the possibility of a coherent representation. For economic and institutional relationships, for example in the case of packaging, the feasibility of detailed representations is demonstrated by works as Sofres (2000). The same applies to technological and material flow representations, as those of RDC- Pira (2003) as those emerging from MFA. The latter can be further developed and detailed following the Wuppertal approach as well as detailed input-output techniques (see Ayres and Ayres 1998 and 1999, Moll et all 2003, Femia et al. 2003). Significant information exists on technical and economic aspects of specific waste treatment technologies, (see Barbiroli and Raggi 1994, EIB 2002) and they are increasingly addressed by LCA approaches.

5. Main results and conclusions

We have examined the issues associated to three general approaches to policy effectiveness analysis that are based on different forms of modelling and empirical estimation techniques. The main results and suggestions are:

1. ***Ex post CBA.*** Systematic experiences of ex ante and ex post CBA applied to policy analysis, typically associated to the US EPA's action under specific regulation requirements, suggest that a rigorous application of this approach is very demanding. CBA is a flexible, well-established, and codified conceptual tool, extensively used at different level of sophistication (different level of complexity, either supported by large-scale models or not, either very aggregated or disaggregated variables, etc.). However, its origin is for ex ante projects evaluation, i.e.

comparison of alternative 'marginal' well-defined changes that do not substantially change the structure of the system. Its use for ex post evaluation of policies can have various limitations given that policy, differently from a 'marginal' project, generally represents, especially in the environmental field, a structural change of the sub-system considered. Furthermore, experiences of application to policy evaluation suggest the need of very huge amount of data that must be all translated in common units, i.e. money. Ex post evaluation with CBA approaches suggest in any case the need of building a countrefactual (i.e. costs and benefits of non-action) for correctly measuring the contribution of policy.

2. ***Experimental methods and counterfactual analysis.*** Conterfactual analysis is the basis of experimental methods increasingly used for appraisal and evaluation of public policies. We have examined the case of labour policies (another area of application is, for example, local development policies). The basic need of building a control sample, to which policy is not applied, to be compared with the sample to which policy is applied in order to derive policy effect, makes this approach suitable for problems in which extensive samples of similar units are available. This is rarely the case with environmental policies where the unit of observations are countries or regions or administrative units that are few in number and are generally all subject to the same policy, or similar policies (in the case, for example, of national policies born from the transposition of EU directives). Nevertheless, there can be areas of environmental policy in which the approach is feasible, e.g. when the units are households or individuals or companies *selectively* subject to regulation. However, at the present state of development of the information base, the areas of applications of experimental methods in environmental policy analysis can be very limited. An equivalent of counterfactual analysis more relevant for environmental policy analysis is the BAU approach (see below).

3. ***Econometric modelling.*** In many fields, including some areas of environmental policy, the rigorous analysis of policy impacts, effectiveness, and efficiency are addressed by using models (be they micro models for social experiments or macro models of the whole economy) that are tractable with statistical and econometric techniques. The use of models is more developed in those areas where there are extensive sets of data and where the complexity of the system is such that "simple" theoretical assumptions about agents' behaviour coupled with massive testing exercises on data sets allow to draw reasonable conclusion on the role of policy in influencing one or more variables addressed by policy as objectives. Where data and theory are not well developed, the problems of appropriate model specification and choice of estimation techniques with limited data tend to prevail. Besides the issue of data availability, we have highlighted some other issues in using econometric models (one equation structural models) for policy evaluation. The first one is the definition of the dependent variable, represented by an (environmental) achievement indicator. It is better to choose an indicator that is as much as possible under the control of the policy considered. Too general or heterogeneous achievement indicators, in terms of factors affecting their levels and changes, can supply misleading suggestions about effectiveness. Preliminary decompositions of the indicators in order to check their direct relations with policy objectives and action can be useful. The second issue is the need of making explicit the indirect relationships among the explanatory variables and in particular their indirect relationship with policy response/action variables. Exogeneity and correlation of explanatory variables must be carefully checked and the structural relationships among variables should be described in details. The third issue is about the definition of policy response/action indicators to be included in the model as explanatory variables. We have considered in particular the different meaning of expenses and costs as policy response variables, the problem of detecting lagged effects of policy variables along a policy-effectiveness cycle, and the issue of representing innovation as a policy-related variable.

We have then examined, also in relations to the above modelling approaches, three specific tools that are often used in making judgements on policy performance and we have tried to highlight their pros and cons in policy effectiveness evaluation. The main results and suggestions are:

1. ***The BAU approach.*** The ex post evaluation approach based on building a Business-as-Usual scenario, estimated without policy action, to which the actual path of the indicator (influenced by policy action) can be compared in search for policy effects, is the equivalent of a counterfactual analysis. In principle, it responds to the requirement of a rigorous analysis. However, we have examined various methodological issues that mostly relate to the way the BAU path, which is not observable, is estimated. Model-based BAU simulations can be generally useful for evaluating the policy-only effects in those cases where policy change is not so pervasive to impair the structural stability of the model (in terms of estimated parameters and exogeneity of explanatory variables). In the case of pervasive policies that change the parameters and the relationships among explanatory variables, the estimated BAU may still be useful but it includes a distortion that depends from being estimated with exogenous variables (after policy introduction) that might be no more exogenous and parameters that might be now different from those before policy. Paradoxically, the usefulness of a BAU estimated as a statistical projection of past values of the indicator before policy, albeit not a good counterfactual, might be considered. If it is a statistical projection, BAU is completely independent from value of exogenous variables after policy and it incorporates only information before polices. Advanced time series techniques, such as ARMA models, can do a good work in exploiting the information incorporated in past data before policy and they might supply a good reference for comparing the after-policy path. The condition is obviously to have a very good availability of data, i.e. long time series. A more favourable position for using models for estimating BAU may emerge when there is already a policy in place and the comparison is between the new and the old policy, if the two are not too radically different. Nevertheless, model-based BAU approaches can be used in some specific cases where simple policy changes are addresses for not-too-complex indicators.

2. ***Comparison of past (before policy) and present (after policy) trends.*** Trends and structural break analysis is not, in general, a rigorous tool for policy evaluation. The comparison of the indicator before and after policy (or with a changing policy) cannot take into account factors different from policy that shaped the changing trend of the indicator. However, for both specific and highly aggregated achievement indicators, the analysis of changing trends and structural breaks can be very useful in reveal whether a policy change has been irrelevant or not. Furthermore, it is conceptually not too different from comparing actual values of the indicator after policy and a BAU if the latter is estimated as a statistical projection of past trends before policy. Many statistical tests are currently available for performing structural break analysis and they can supply a sound technical support for claming (or not) that the introduction of a policy or its change corresponds to a structural change on the indicator representing the achievement of policy objective. Obviously, even in the case that breaks in the indicators time series soundly correspond to policy introduction, we rest with the task of more complete analysis of the actual role of policy.

3. ***Distance from target analysis.*** It is apparently a simple way for defining is we are converging or diverging with respect to a policy-determined path and then if our policy implementation is effective. However, the interpretation of where we are and where we should be for defining our policy as effective, is not without ambiguities. It remains in any case an implicit and non-explicative evaluation of effectiveness. Econometric models can reduce the ambiguities by suggesting, for example, where we should be at a certain point of time according to a convergent path, or what level of policy actions we should implement. The cost is again that of good model building and good data availability. By exploiting the suggestions of economic growth studies about convergence and non-convergence, and given the limited availability of time series, a possible different use of the distance-from-target approach can be tried by

performing cross-countries comparison and benchmarking for a specific policy target. Leaving aside the possible risk of such an analysis, it could allow, in principle, the comparison of the performance of different policy packages, instead of single instruments, adopted by different countries. The analysis of the role instruments had for the different performance can be a second step to be dealt with different approaches, including those considered above. Cross-countries benchmarking can be feasible even with a limited set of data and can be easily integrated in qualitative analysis of policy performance in different countries or regions.

The main suggestions about the use of structured methodologies for policy effectiveness analysis are therefore the following:

- The use of *ex post CBA* cannot be excluded as a useful instrument for evaluation. In principle, it is an ideal approach for measuring global welfare implications of a policy, and it is extensively used at different levels. However, there are possible conceptual objections to ex post policy evaluation based on CBA if the policy departs from the condition of being a marginal change. Furthermore, there can be strong data limitations, also in terms of quality, especially for cost data, that prevent from performing good CBA of many environmental areas. Its use should therefore be limited to the analysis of small and well-defined environmental policies, or even better policy changes, for which there is very sound information.
- The potential use of *experimental approaches* seems to be very limited for environmental policy analysis. It is feasible for very specific policies addressing a large number of units, be they households, local administrations, or companies.
- *Econometric modelling* is a flexible instrument for policy evaluation. However, good modelling is usually possible when a cumulative work of modelling and information-base building has been made during time, which is not the case with all environment-related sectors. Intensive variables testing in search for statistical significance of candidate explanatory variables is not possible for many environmental sectors. For making econometric modelling a useful approach to policy evaluation there is the need to build up detailed representations of the systems addressed by policies and, at the same time, more extensive collections of socio-economic data. Nevertheless, models are a feasible instrument for some environmental policy areas.
- The use of model-based *BAU analysis* is a feasible approach provided that it is used with non-systemic policies and then with well-defined policy changes that do not change the basic structure of the system addressed. Advanced statistical techniques can supply good estimates of BAU without models, provided that long time series of data are available.
- The *analysis of past and after-policy trends and structural breaks* is ways for identify the existence of changes reliably attributable to policy introduction or change. After establishing that policy mattered for the achievement indicator, the issue of more interpretations about why and how remains open.
- The use of *distance from targets* can be usefully explored in a cross country or cross-regional setting in which, by exploiting convergence analysis, benchmarking analysis can be done. Even in this case, the interpretation of why and how a country is performing better than others in terms of target achievement for a specific policy requires modelling or mixed qualitative and quantitative analysis.

A very general conclusion can be that all quantitative techniques have some limitations and policy effectiveness evaluation cannot disregard the information coming form qualitative analysis of policy effectiveness. The latter can also supply very useful information for models building and specification.

As a final point, we have suggested a research direction aiming at developing detailed system representations as a basis of good and non-ambiguous definition of how policy works and through which channels it arrives to influence the indicators of target/objective. The basic consideration is that systems of indicators such as DPSIR have intrinsically a causative structure in relating different types and areas of single indicators. The detailed representation of both the socio-economic system and the policy process is therefore a necessary step in clarifying the causative relationships between, on the one hand, policy Responses and, on the other hand, Determinants and Pressures.

References

Arrow, K., et al, 1996, Benefit Cost Analysis in Environmental, Health and Safety Regulation, American enterprise Institute, The Annapolis Center and Resource for the Future, Washington D.C.

Ayres R., Ayres L.W., 1998, *Economy-wide Applications of Mass-balance Principles to Materials and Waste*, Edward Elgar, Cheltenham.

Ayres R., Ayres L.W., 1999, *The Life Cycle of Materials*, Edward Elgar, Cheltenham.

Biondi V., Zoboli R., Mazzanti M, et al., 2002, *Nuove prospettive del Protocollo di Kyoto: meccanismi attuativi ed impatto sulla competitività*, IPA e Ministero dell'Ambiente, Milano.

Bodansky D., 2002, Implications for US companies of Kyoto's entry into force without the United States, PEW Center on Global Climate Change, Arlington.

Borghesi, S., Vercelli, A., 2003, Sustainable Globalisation, *Ecological Economics*, vol.44, pp.77-89.

Boyd, J., 2003, Water Pollution Taxes: a good idea doomed to failure?, discussion paper 03/20 Resources for the Future, Washington D.C.

Bressers, H., 1988, A Comparison of the Effectiveness of incentives and Directives: the case of Dutch Water Quality Policy, *Policy Studies Review*, n.3, pp. 500-518.

E3M Lab, Capros P., Mantzos L., 2000, The Economic Effects of EU-Wide Industry-Level Emission Trading to Reduce Greenhouse Gases, Athens, May.

EEA, 1997, Environmental Agreements. Environmental Effectiveness, European Environment Agency, Copenhagen.

EEA, 2001, Reporting on environmental measures: Are we being effective?, Environmental Issue Report No. 25, Copenhagen, November.

EEA, 2002, Environmental signals 2001, Copenhagen.

EIB, 2002, EIB financing of solid waste management projects. Evaluation report, EIB, Luxembourg.

EPA, 1997, The Benefits and Costs of the Clean Air Act: 1970-1990, Office of Policy Analysis, Washington D.C.

EPA, 1990, Environmental Investments: the Costs of Clean Environment, Office of Policy Analysis, Washington D.C.

ETC-ACC, 2003, Response indicators for air and climate change, draft note, January.

European Commission, 2000, A Study on Investment and Employment related to EU Policy on Air, Water and Waste. Best estimates of Costs, European Commission, Brussels

European Commission, 1996, Cost-Benefit Analysis of the Different Municipal Solid Waste Management Systems: Objectives and Instruments for the Year 2000, Final Report, March, Brussels.

European Commission, 2000, European Climate Change Programme. Progress Report, Brussels, November.

Farrow, S., Toman, M., 1998, Using Environmental Benefit Cost Analysis to Improve Government Performance, Resource for the Future, discussion paper 99/11, Washington DC

Freeman, M., 2002, Environmental Policy since Earth day I: what have we gained?, *Journal of Economic Perspectives*, vol.16, n.1, pp.125-146.

Frondel, M., Schmidt, C., 2001, Evaluating Environmental Programs. The Perspective of Modern Evaluation Research, Discussion paper 397, IZA, Bonn.

Hahn, R.W., 1998, Government analysis of the benefits and costs of regulation, *Journal of Economic Perspectives*, vol.12, n.4, pp.201-210.

Hemmelskamp J., 1997, Environmental Policy Instruments and their Effects on Innovation, *European Planning Studies*, Vol. 5, No. 2.

Hemmelskamp J., Leone F., 1998, The Impact of EU Regulation on Innovation of European Industry, IPTS-JRC, EUR 18111 EN, Seville.

Jaffe A.B., Newell R.G., Stavins R.N., 2003, Technological Change and the Environment, in: Mäler K-G. and Vincent J.R. (eds.), *Handbook of Environmental Economics*, Volume 1, Elsevier, Amsterdam.
Jenkins, R., Martinez, S., Palmer, K., Podolsky, M., (2003), The Determinants of Household Recycling: a Material Specific Analysis of Recycling Program Features and Unit Pricing, *Journal of Environmental Economics and Management*, vol. 45, pp.249-318.
Kemp R., 1997, Environmental Policy and Technical Change, Edward Elgar, Cheltenham.
Klemmer P. (eds.), 1999, *Innovation and the Environment*, Analytica, Berlin.
Mazzanti M., Simeone M.G., Zoboli R., 2003, Policy response indicators for packaging waste policy, ETC-WMF, WP2j, Copenhagen, December.
Mazzanti M., Zoboli R., 2005, Environmental Policy Instruments and Induced Innovation: the case of European end of life vehicles Directive, *Ecological Economics*, forthcoming
Morgenstern, R., 1997, Economic Analysis at EPA: Assessing Regulatory Impact, RFF, Washington DC
Nordhaus, W.D., 1992, Lethal Model 2: The Limits to Growth Revisited, Brooking Papers on Economic Activity, No. 2.
Palmer K., Sigman, H., Walls M., 1996, The Cost of Reducing Municipal Solid Waste, Discussion Paper 96-35, Resources for the Future, September.
Quadrio Curzio A., Zoboli R., 1996, The cost of sustainability, Proceedings of the conference "The 50th Anniversary of the United Nations and the Italian contribution towards the realisation of the Earth Charter", Italian National Academy of Sciences, Rome.
RDC Environment & PIRA International, 2003, Evaluation of costs and benefits for the achievement of reuse and recycling targets for the different packaging materials in the frame of the packaging and packaging waste directive 94/62/EC, Final consolidated Report.
Risk and Policy Analysis Ltd, 1998, Economic evaluation of environmental policy and legislation, Final Report to the European Commission, DGIII, Norfolk.
Simeone M.G., Zoboli R., 2002, Policy effectiveness/response indicators: a suggested approach, APAPT and ETC-WMF, mimeo, July.
Taylor, Nelson Sofres Consulting, 2000, Cost-efficiency of packaging recovering system. The case of France, Germany, the Netherlands and the United Kingdom, Report to DG Enterprise, Commission of the European Communities, February.
VV.AA, 1998, Getting Environmental Policy Right. The Rational Design of European Environmental Policy from an Economic Perspective, European Economy, n. 1, Brussels.
Weyant J.P, 1993, "Costs of reducing global carbon emissions", *Journal of Economic Perspectives*, Vol. 7, N. 4.
Weyant J.P., 2000, An Introduction to the Economics of Climate Change Policy, Pew Center of Climate Change, Washington DC.
Zoboli R., Barbiroli G., Leoncini R., Mazzanti M., Montresor S., 2000, Regulation and Innovation in the Area of End-of-Life Vehicles, Final Research Report, edited by Fabio Leone, IPTS-JRC, EUR 19598 EN, March, Seville.

Other works examined

Anderson M.S., 1999, Governance by green taxes: implementing clean water policies in Europe 1970-1990, *Economics and Policy Studies*.
Baumol W., Oates W, 1988, *The Theory of Environmental Policy*, Cambridge, Cambridge University Press.
Bohm P., Russel C., 1985, Comparative analysis of alternative policy instruments, in Kneese A.V. and Sweeney J.L., *Handbook of Natural Resource and Energy Economics*, Vol I, North Holland, Amsterdam.
Brunnermeier, S., Cohen, M., 2003, Determinants of Environmental Innovation in US Manufacturing Industries, *Journal of Environmental Economics and Management*, vol. 45, pp. 278-93.
Cabinet Office, 1998, A Guide to Regulatory Impact Assessment, London
Chang H-J., 1997, The Economics and Politics of Regulation, Cambridge Journal of Economics, No. 21.
Congleton, R. (ed.), 1996, *The Political Economy of Environmental Regulation*, University of Michigan Press, Ann Arbor
Cropper M.L. and Oates W.E., 1992, Environmental Economics: A Survey, *Journal of Economic Literature*, Vol. XXX, June.
CSERGE, EFTEC and Warren Spring Laboratory, 1993, Externalities from Landfill and Incineration, CSERGE, UCL, London

DCMS, 2000, "The White Book" Option Appraisal of Expenditure Decisions: A Guide for DCMS and its Sponsored Bodies, DCMS Technical Paper No. 2, DCMS, London

Coopers Lybrand, CSERGE e EFTEC, 1996, Cost Benefit Analysis of the Different Municipal Solid Waste Management Systems: Objectives and Instruments for the Year 2000, Luxembourg: European Commission

Dean, T.J., Brown, R.L., Stango, V., 2000, Environmental Regulations as a Barrier to the Formation of Small Manufacturing Establishments: a Longitudinal Analysis, *Journal of Environmental Economics and Management*, vol. 40, pp. 56-75

DEFRA, 2000, Waste Strategy 2000 for England and Wales, part I & II, London

DEFRA, 2001, Implementation of Council Directive 1999/31/EC on the Landfill of Waste, Second consultation paper, London

DETR, 1999, A Way With Waste: a Draft Waste Strategy for England and Wales, Parts One and Two, London, DETR

DETR, 1999, Limiting Landfill: a Consultation Paper on Limiting Landfill to Meet the EC Landfill Directive's Targets for the Landfill of Biodegradable Municipal Waste, London: DETR

DETR, 2001, Tradable Landfill Permits: Consultation Paper, London: DETR

DETR. 2000. Waste Strategy 2000 for England and Wales, Parts One and Two, London: DETR

Dole, D., 2001, Measuring the Impact of Regulation on Small Firms, National Centre for Environmental Economics, EPA, www.epa.gov

EEA, 2003, Europe's environment: The third assessment, Environmental assessment report No. 10, Copenhagen.

ERM, 2001, Research Study on International Recycling Experience, London: DETR.

Hahn R., Stavins R.W., 1991, Economic incentives for environmental protection: integrating theory and practice, Discussion Paper QE92-08, Resources for the Future, Washington D.C.

Harrington W., Morgensten R.D., Nelson P., 1999, On the accuracy of regulatory cost estimates, Discussion Paper 99-18, RFF, Washington D.C.

Hitchens D. et alii, 2001, The impact of best available techniques (BAT) on the competitiveness of the European industry, IPTS-JRC, November, Seville.

Hitchens D.M., 1999, The implications for competitiveness of environmental regulation for peripheral regions in the EU, Omega, *International Journal of Management Sciences*, 27.

IPTS-JRC, 1997, Environment and Employment, Institute for Perspective Technological Studies, Joint Research Centre, European Commission, Seville.

Jacobs, M., 1995, Do Environmental Policies Cost Jobs?, Employment Policy Institute, Economic Report, Vol. 9, No. 8.

Jaffe A.B., Peterson S.R., Portney P.R., Stavins R.N., 1995, Environmental regulation and the competitiveness of US manufacturing: what does the evidence tell us?, *Journal of Economic Literature*, vol. 33, March.

Kremers, H., Nijkamp, P., Rietveld, P., 1999, The Scope of Meta-analysis for Transport Policy Impact Analysis in Environmental Economics, Series Research Memoranda, n.34, 1999, Vrije Universiteit, Amsterdam.

Mazzanti M., Zoboli R., 2001, "Environmental policy instruments, cost-minimising strategies, and induced innovation: the case of end-of-life vehicles", proceedings of the 2001 Environmental Policy and the Costs of Compliance Research Workshop, September 18th, 2001, ERP Environment, London School of Economics, pp. 138-151.

Mickwitz, P., 2003, Is it as bad as it sounds or as good as it looks? Experiences of Finnish Water Discharge Limits, *Ecological Economics*, forthcoming.

Milliman S.R. and Price R., 1989, Firm incentives to promote technological change in pollution control, *Journal of Environmental Economics and Management*, vol. 17, n.3.

Natale P., 1994, An Economic Analysis of the Toepfer Law, in Quadrio Curzio, Prosperetti, Zoboli (Eds.), *The Management of Municipal Solid Waste in Europe*, Elsevier Science Publishers, Amsterdam.

OECD, 1997, Evaluating Economic Instruments for Environmental Policy, OECD, Paris.

OECD, 2001, Environment related taxes in OECD countries: Issues and strategies, OECD, Paris.

Palmer, K., Walls, M., 1999, Extended Producer Responsibility: an Economic Assessment of Alternative Policies, Resource for the Future, discussion paper 99/12, Washington DC

Salmons, R., 2002, New Areas for Application of Tradable Permits Solid Waste Management, report produced for the National Policies Division, OECD Environment Directorate, CSERGE/UCL, London

Simeone M.G., Moll S., Brodersen J., Juul J., Mortensen L., Visiimaa M., Zoboli R., 2001, Development of an Indicator Framework on Waste and Material Flow, Technical Report, ETC/WMF, December.

Symonds, ARGUS, COWI, PRC Bouwcentrum, 1999, Construction and demolition waste management practices, and their economic impacts, Report to DGXI, European Commission, February

Treasury, 1997, Appraisal and Evaluation in central Government ("The Treasury Green Book"), London

Wiener J.B., 2000, Policy Design for International Greenhouse Gas Control, Resources for the Future, Washington DC.

Zoboli R., 1994, "The Integrated Use of Economic Instruments in the Policy of Municipal Solid Waste", in A. Quadrio Curzio, L. Prosperetti, R. Zoboli, (eds.), The Management of Municipal Solid Waste in Europe: Economic, Technological and Environmental Perspectives, Elsevier Science Publishers Plc., Amsterdam, pp. 234-244.

Zoboli R., 2000, Environmental Regulation and Innovation in the End-of-Life Vehicles Sector, in Hellemskamp J., Rennings K., Leone F. (eds.), *Innovation-Oriented Environmental Regulation*, ZEW Economic Studies 10, Physica-Verlag, Heidelberg.

Zoboli, R., 2000, Valutazione delle politiche ambientali e sistemi di indicatori, in Lupò Avagliano M.V. (a cura di), *Efficienza della pubblica amministrazione*, Franco Angeli, Milano.