L'analisi economica del diritto ambientale

A CURA DI MICHELE CARDUCCI

Working Paper del Laboratorio di analisi ecologica del diritto
QUADERNO N. 8

INDICE DELL'ANTOLOGIA

BATEMAN I.J. – MACE G.M. – FEZZI C. – ATKINSON G. – TURNER K.
Economic Analysis for Ecosystem Service Assessments

DI GIULIO E.
Analisi economica del diritto e ambiente

FAURE M.G.
Economic Analysis of environmental Law: An Introduction

GHOSH J.
Economic Analysis of Environmental Laws and Regulations

HONGJUN C. – HAOHAO S. – XUESEN C.
Analysis of Environmental Law Enforcement Mechanism Based on Economic Principle

NISHA J. – SHEN S.
A study on the relationship of environmental regulations and economic performances

NUNES P.A.K.D. – VAN DEN BERGH J.C.J.M. – NIJKAMP P.
Ecological–Economic Analysis and Valuation of Biodiversity

PITEA C.
Right to Property, Investments, and Environmental Protection: The Perspectives of the European and Inter-American Courts of Human Rights

POLASKYA S. – KLING C.L. – LEVIN S.A. – CARPENTER S.R. *ET AL.*
Role of economics in analyzing the environment and sustainable development

ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/227451208

Economic Analysis for Ecosystem Service Assessments

Article *in* Environmental and Resource Economics · January 2011

Source: RePEc

CITATIONS
160

READS
1,074

5 authors, including:

Ian J. Bateman
University of Exeter
370 PUBLICATIONS **21,392** CITATIONS
SEE PROFILE

Georgina M Mace
University College London
187 PUBLICATIONS **16,417** CITATIONS
SEE PROFILE

Giles Atkinson
The London School of Economics and Political Science
91 PUBLICATIONS **6,679** CITATIONS
SEE PROFILE

R. Kerry Turner
University of East Anglia
340 PUBLICATIONS **28,936** CITATIONS
SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Project Lancet Countdown: Tracking the connections between health and climate change View project

Project Valuing the Arc View project

All content following this page was uploaded by Ian J. Bateman on 05 June 2014.

The user has requested enhancement of the downloaded file.

Environ Resource Econ (2011) 48:177–218
DOI 10.1007/s10640-010-9418-x

Economic Analysis for Ecosystem Service Assessments

Ian J. Bateman · Georgina M. Mace · Carlo Fezzi · Giles Atkinson · Kerry Turner

Accepted: 2 May 2010 / Published online: 13 October 2010
© The Author(s) 2010. This article is published with open access at Springerlink.com

Abstract The paper seeks to contribute to the expanding literature on ecosystem service assessment by considering its integration with economic analyses of such services. Focussing upon analyses for future orientated policy and decision making, we initially consider a single period during which ecological stocks are maintained at sustainable levels. The flow of ecosystems services and their contribution to welfare bearing goods is considered and

We are grateful for comments from two anonymous referees, Sir Partha Dasgupta (Cambridge) and Karl-Göran Mäler (Beijer Institute) of the UK NEA Economics Advisory Board, Brett Day (UEA), Ed Barbier (Wyoming), Grischa Perino (UEA), Kirk Hamilton (World Bank), Jules Pretty (Essex) and all members of the wider UK NEA economics team, especially Nicola Beaumont (Plymouth); Katherine Bolt (RSPB); Brendan Fisher (Princeton); Nick Hanley (Stirling); Andreas Kontoleon (Cambridge); Pushpam Kumar (Liverpool); David Maddison (Birmingham); Dominic Moran (SAC); Joe Morris (Cranfield); Susana Mourato (LSE); Unai Pascual (Cambridge); Steve Polasky (Minnesota); Daan van Soest (Tilburg & VU Amsterdam); Gregory Valatin (Forestry Research, Scotland) and Bhaskar Vira (Cambridge). Opinions and remaining errors are the sole responsibility of the authors. The UK NEA is co-chaired by Defra's Chief Scientific Adviser, Professor Robert Watson (previous Chair of the Millennium Ecosystem Assessment), and by Professor Steve Albon of the Macaulay Institute. UNEP-WCMC are the UK NEA secretariat from which we are grateful to Claire Brown for comments and Lucy Simpson and Megan Tierney for support.

I. J. Bateman (✉) · C. Fezzi · K. Turner
Centre for Social and Economic Research on the Global Environment (CSERGE),
School of Environmental Sciences, University of East Anglia (UEA), Norwich NR4 7TJ, UK
e-mail: i.bateman@uea.ac.uk

I. J. Bateman
University of Western Australia, Perth, Australia

I. J. Bateman
University of Waikato, Hamilton, New Zealand

G. M. Mace
Centre for Population Biology, Imperial College London, Silwood Park, Ascot, UK

G. Atkinson
Department of Geography and Environment and Grantham Research Institute on Climate Change and Environment, London School of Economics and Political Science, London, UK

methods for valuing resultant benefits are reviewed and illustrated via a case study of land use change. We then broaden our time horizon to discuss the treatment of future costs and benefits. Finally we relax our sustainability assumption and consider economic approaches to the incorporation of depleting ecological assets with a particular focus upon stocks which exhibit thresholds below which restoration is compromised.

Keywords Ecosystem services · Economic valuation · Stocks and flows · Wellbeing · Spatial and temporal issues · Sustainability · Irreversibility · UK National Ecosystem Assessment (UK NEA)

JEL Classification Q01 · Q10 · Q20 · Q30 · Q51 · Q56 · Q57

1 Introduction

The crucial role which natural systems play in underpinning economic activity and anthropocentric wellbeing is of growing concern as evidence mounts of the increasing pressures being placed upon such systems by human activity (GEF 1998; Chapin et al. 2000; Koziell 2001; MA 2005; CBD 2006; Loreau et al. 2006). One reflection of that concern is the recent undertaking of major global assessments of the status of the services provided by ecosystems (see, for example, MA 2005 or TEEB 2009). Economic analysis is an increasing feature of such undertakings and has prompted a rapidly expanding literature regarding the implementation of such analyses (see, for example, Bockstael et al. 2000; Balmford et al. 2002; de Groot et al. 2002; Howarth and Farber 2002; Heal et al. 2005; Barbier 2007; Boyd and Banzhaf 2007; Wallace 2007; Fisher and Turner 2008; Fisher et al. 2008; Mäler et al. 2008; Tschirhart 2009; Liu et al. 2010; Turner et al. 2010). It is the intention of the present paper to contribute to this literature in a number of ways. In particular we draw upon this preceding literature to propose a general framework and nomenclature for integrating economic analyses within ecosystem service assessments. Given the inherently interdisciplinary nature of such undertakings, we review some of the fundamental principles of economic analysis so as to introduce these to a natural science audience. At the same time we attempt to address certain key challenges which economists will have to face in order to adequately represent the complex nature of ecosystem service provision within economic analyses.

This paper has a direct empirical context as it provides much of the economic methodology underpinning the ongoing UK National Ecosystem Assessment (UK NEA). This exercise, coordinated by UNEP-WCMC, is a direct successor to the UN Millennium Ecosystem Assessment (MA 2005), adapted for the specific context of the UK and taking on a number of conceptual and scientific advances which have arisen since 2005. The MA conceptual framework linked primary 'supporting services' (e.g. soil formation) to 'provisioning' (e.g. food production), 'regulating' (e.g. climate) and 'cultural' (e.g. recreation) services. Its findings not only demonstrated the importance of ecosystem services to human wellbeing, but also showed that at global scales, many key services are being degraded and used unsustainably. The MA challenge was taken up by the UK House of Commons Environmental Audit Committee (2007) which recommended that, 'ultimately the Government should conduct a full MA-type assessment for the UK to enable the identification and development of effective policy responses to ecosystem service degradation'. Throughout this paper we illustrate the principles put forward by reference to the empirical approach adopted by the UK NEA.[1]

[1] For further details see http://uknea.unep-wcmc.org/.

This is not a conventional economics journal paper. Rather it is intended as a means of introducing both economists and non-economists (and in particular natural scientists) to the UK NEA and through that to the wider principles involved in the application of economic analysis techniques to ecosystem service assessments.[2] As such there are certain sections which review basic economic principles. We hope that the regular audience of this journal will tolerate this and find something of interest in the application of such techniques to the issue of ecosystem services.

Ecosystem service assessments and accompanying economic analyses can be roughly divided into two types.[3] 'Sustainability analyses' are typically focused upon changes occurring up to the present day and involve assessing whether or not the observed development path of an economy has been progressively running down the natural asset base which underpins its viability. Such analyses provide important early warning signals for non-sustainable growth patterns (see for example, MA 2005 and TEEB 2009). However, once such warnings have been raised 'programme evaluation' analyses provide forward-looking, policy relevant assessments and comparisons of available alternative future development strategies. Such analyses are typically implemented through assessments of one or more future scenarios, the drivers of which include forecasts of environmental change, trends in domestic and world markets and potential (and possibly dynamically compensating) policy shifts. The UK NEA undertakes both forms of analysis, encompassing both a sustainability analysis of historic trends from the middle of the last century to the present day, as well as programme evaluations of different, alternative futures from the present until 2060 as captured in a series of scenario analyses. Given the prior literature on sustainability evaluations, the present paper focuses upon the forward looking programme evaluation element of the UK NEA and in particular the economic analyses to be conducted as part of this.

True to the economics roots of this paper we start by making a number of assumptions. However, we subsequently relax these to arrive at what we hope is a generally applicable approach to undertaking economic analyses for policy orientated, forward looking, ecosystem assessments of alternative scenarios.

Our initial assumptions hold constant two complicating factors; the passage of time and the sustainability of stocks of ecosystem assets. We neutralise the former issue by initially focussing upon a given scenario operating in a single period across which we assume individuals are indifferent as to when flows are received.[4] Considering the latter issue and noting that ecosystem services are flows generated by stocks of ecosystem assets[5] (e.g. the flow of trees for timber is generated by the stock of global forests), we further simplify by initially assuming that the rate at which ecosystem services are extracted is sustainable in that it does not reduce stocks over the assessment period (e.g. harvesting of timber stays within the range of natural regeneration). We relax both assumptions subsequently, but with these in place we now consider how one might conduct an economic analysis of ecosystem service flows across the period.

We start by discussing the overall conceptual framework of our approach and developing an accompanying terminology. We use this to consider the web of natural world relations underpinning each ecosystem service and acknowledge the role which human and manufactured capital play in combining with such services in the production of welfare bearing goods (e.g. trees are an insufficient input to generate timber which also requires human skills

[2] We are grateful to the Guest Editors for the invitation to submit such an unconventional paper.

[3] We are grateful to Sir Partha Dasgupta for highlighting this distinction.

[4] Whether such a period actually exists we do not debate as we relax this assumption subsequently.

[5] We are grateful to Karl-Göran Mäler for his advice and invaluable comments regarding this issue.

and technology).[6] We then consider methods for isolating the ecosystem service element within the production of such goods before considering the various methods used to value ecosystem services. These principles are put into practice through a related series of illustrations considering the issue of land use. These examples also serve to underline the need to ensure that economic analyses recognise the inherently spatial nature of most ecosystem services. This discussion is also used to relax our single period assumption. By allowing the benefits and costs of ecosystem services to occur across multiple periods we highlight the treatment of time within economics through the process of discounting. Finally we relax our sustainability assumption and consider cases where the 'harvesting' of ecosystem service flows may lead to the long term depletion of natural assets. We conclude our discussions by considering alternative approaches to dealing with such problems.

2 Conceptual Framework and Terminology

The Millennium Ecosystem Assessment defines ecosystem services as "the benefits people obtain from ecosystems" (MA 2005; p. 53). Fisher and Turner (2008) expand on this definition to propose that "ecosystem services are the aspects of ecosystems utilized (actively or passively) to produce human well-being" (p. 2051). Both definitions clarify the anthropocentric focus of the ecosystem service concept.[7] While a wider understanding of environmental processes may be a necessary part of any ecosystem assessment, it is the role of the natural world in delivering human wellbeing which is central to assessments of ecosystem services. It is this human focus that necessitates the integration of economic analysis within such assessments so that we can quantify and value ecosystem services ensuring that their importance and worth can be incorporated within decision making.

As mentioned previously, the level of ecosystem service 'harvested' within any given period can be thought of as a 'flow' extracted from an underlying 'stock' of ecosystem asset (Barbier 2009; Mäler et al. 2009).[8] Just as with a stock of financial wealth in a bank, the withdrawal rate can either be sustainable (say an amount which is less than or equal to the real value of interest paid in that period) or unsustainable (an amount which, if maintained, will eventually deplete the real value of the asset to levels which then reduce the available flow of income).[9] However, we initially consider just a single period throughout which the

[6] Changes in scenario are reflected by alterations in the levels of services and other inputs resulting in altered outputs of goods.

[7] This also alerts us to the potential conflicts which may arise when decision making is based on preferences. Some people may consider certain species' repulsive or dislike entire ecosystems. In a recent economic valuation study, respondents gave positive values for reductions in the extent of coastal mudflats (Bateman et al. 2009b); a finding which reflected the negative visual amenity of such areas as perceived by many people. Of course these need to be set against the biodiversity habitat, coastal defence, pollution regulation and other benefits such areas provide (Barbier 1994). Nevertheless decision making through economic valuation is a reflection of human preferences. This is, with some caveats (certain of which we discuss subsequently), in line with a democratic ethos but as such exposes assessments to the tyranny of the majority. At the same time it shows that pressure groups can attempt to influence the outcome of economic analyses of ecosystem services by changing peoples' preferences.

[8] Having a larger *stock* of ecological assets might mean that more services will be enjoyed although, as Barbier et al. (2008) and Boyd and Krupnick (2009) note, the relationship may be non-linear and lagged.

[9] Note that economists will sometimes refer to flows as income and to stocks as wealth, the true intergenerational value of which is referred to as 'inclusive' or 'comprehensive' wealth (Arrow et al. 2007; Dasgupta 2009; Mäler 2008; Mäler et al. 2009) as discussed subsequently.

rate of flow extraction is sustainable in that it does not deplete stocks (an assumption which we relax subsequently).[10]

From an economic perspective then, ecosystem services are those contributions of the natural world which generate goods which people value. However, this statement requires some qualification. First, by 'goods' we mean any item or construct that increases human welfare.[11] This includes physical products (e.g. the role of ecosystem services in the production of food) and less tangible goods (e.g. detoxification services). It also includes items which generate use values (such as those just mentioned) and non-use goods which are valued purely for their continued existence (e.g. some unobserved biodiversity). Second, although the valuation of such goods relies upon standard economic theory (Daily 1997; MA 2003; Pagiola et al. 2004; Heal et al. 2005; Barbier 2007, 2009; Sukhdev 2008), these values might not be perfectly or even partially reflected in market prices and therefore require the application of non-market valuation techniques (subsequently we briefly discuss goods which may not be amenable to such methods). This process of uncovering the true value of goods and using this to ensure decisions contribute to improving human welfare is the defining rationale for economic analysis[12] and can result in major gains in wellbeing even when (as currently assumed) the use of ecological assets is sustainable; it obviously takes on an even more crucial role in cases (as considered subsequently) of unsustainable ecological exploitation.

As noted in our introduction, the focus of this paper is to consider the role of economic analysis for ecosystem service assessments within the context of supporting future orientated policy decision making at a project or programme level. This emphasis imposes some structure upon both the economic and natural science elements of ecosystem service assessments. In particular, Fisher et al. (2008) note the potential for double counting errors if an attempt is made to directly value those ecological processes (e.g. weathering, soil formation, nutrient cycling, etc.) which support multiple ecosystem services. The concern here is that if we value both these primary ecological processes (functions) and what we can term the 'final ecosystem services' which directly generate wellbeing or directly contribute to the production of goods, then we are liable to overestimate the total values generated. An obvious concern in adopting such an approach is that an over-concentration on final ecosystem services and their use and non-use to humans may place underlying ecological assets at risk (Gren et al. 1994; Turner et al. 1999) thereby risking over exploitation and system change or collapse. This has to be guarded against by imposing the constraint that ecosystem assets are not run down to unsustainable levels; an issue we are currently assuming away and which we return to focus upon subsequently.

[10] Of course some resources are physically non-renewable, for example coal stocks. This highlights the fact that we are looking at the maintenance of services rather than the physical constitution of any given asset. So we might run down the stock of coal yet maintain the service of energy provision by increasing stocks of alternative energy resources. However, other resources, such as global oxygen stocks, are effectively non-renewable and non-substitutable (Ayres 2007).

[11] This definition is therefore a deliberate conflation of the three forms of ecosystem service benefit defined by Daily (1997) and Barbier (2007), namely: (i) "goods" (e.g., products obtained from ecosystems, such as resource harvests, water and genetic material), (ii) "services" (e.g., recreational and tourism benefits or certain ecological regulatory functions, such as water purification, climate regulation, and erosion control), and (iii) cultural benefits (e.g., spiritual and religious beliefs, heritage values).

[12] This is very different from accountancy with which economics is sometimes erroneously confused. The accountant focuses upon the market price of goods (which for many environmental goods is zero), while the economist is interested in their true value.

We can now define a set of terms for subsequent use and which are intended to be intuitively accessible to both natural scientists and economists:[13]

- We place ecosystems in their particular spatially and temporally defined context. This requires us to know the relevant boundary conditions (e.g. elevation, slope, climate, etc) which define this context at any given location.
- Given these boundary conditions we can observe the pertaining ecosystem structure (e.g. animals, birds, plants and their connections, etc.) and processes (e.g. nutrient cycling etc).
- 'Ecological assets' are the stocks of potential services which the ecosystem, conditioned by structure and processes, might provide. In economic terms we can think of these as the 'wealth' of the ecosystem.
- We use the term 'ecosystem services' to refer to the flow of services (outcomes of structure and processes) provided by ecological assets in some assessment period.
- 'Final ecosystem services' are simply the last item in the chain of ecosystem functioning which inputs to the production of goods. They are the aspects of the natural environment which most directly affect human wellbeing during an assessment period and can also be thought of as the natural 'income' enjoyed during that time. The focus on the final item in the chain of ecosystem services is simply to avoid the double counting which would occur if we also included those more primary and intermediate supporting services. However, while a focus upon final ecosystem services significantly simplifies the economic valuation task, it is insufficient to ensure sustainability and we return to consider this latter problem later in this paper.
- 'Good' is the term we use for any object or construct which generates human wellbeing.[14] This includes both physical and non-physical (pure experiential) objects. So a beautiful natural landscape generates amenity views which are a good to the outdoor walker as much as a piece of timber is a good to the home improver. As noted, some of these items come straight from the natural world without the intervention of humans; the visual amenity of the natural landscape being an example of this (here the final ecosystem service and the good are identical). In contrast other items (like our timber example) require some inputs of manufactured or other human capital.
- 'Benefit' is simply the change in human wellbeing generated by a good. Previous ecosystem service assessments have on occasion used the terms 'good' and 'benefit' synonymously. However, we draw a sharp distinction to highlight the fact that the same good can generate very different benefit values depending on its context (e.g. location) and timing of delivery.[15] Note that some goods generate instrumental 'use value' (e.g. the timber example), while others deliver 'non-use value' (e.g. the knowledge that biodiversity is being conserved even if the person expressing the value does not observe the species concerned). These benefits include both use and non-use values.

[13] We acknowledge that these terms differ from some of those which economists have used previously; see for example Heal et al. (2005).

[14] Note also that this definition of good embraces the economic definitions of both goods and economic services (the latter term including non-consumptive, experiential items). This is a deliberate simplification intended to enhance understanding between the natural and social sciences and avoid a double meaning of the word 'service' being used.

[15] For example, considering the spatial context of a good, a woodland situated on the edge of a major city will generate much greater recreation benefits than a physically identical woodland situated in some remote area. Of course biodiversity might be inversely related to urban proximity. Analysing such trade-offs are the essence of environmental economics.

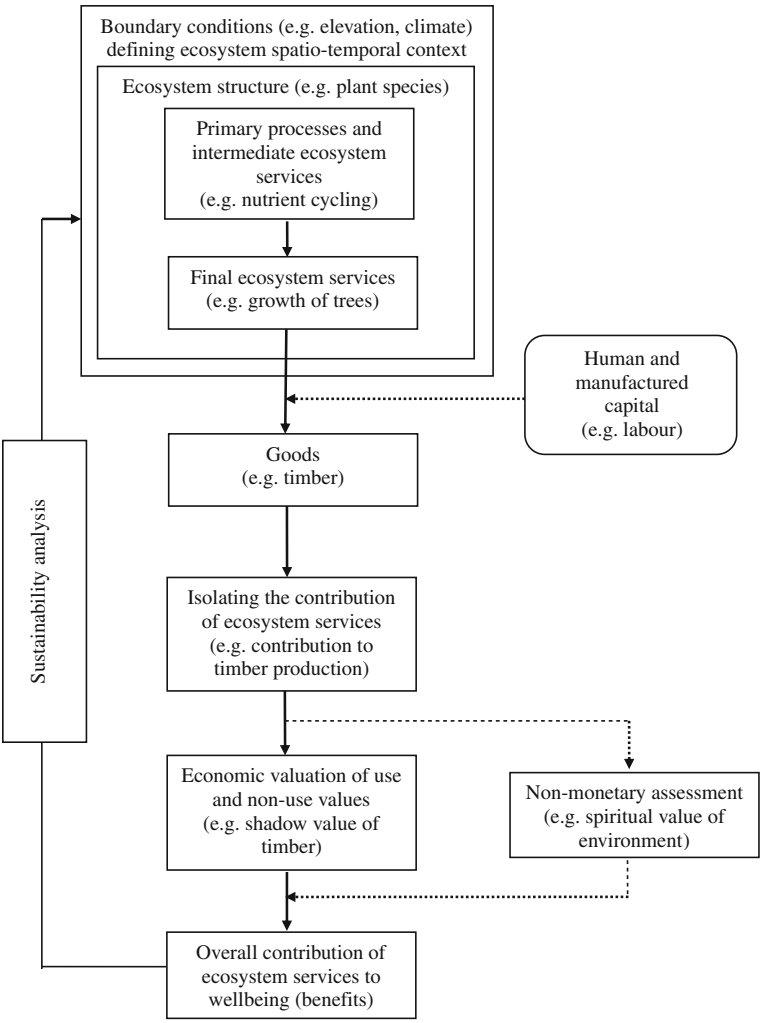


Fig. 1 Phases of a joint ecosystem assessment and economic analysis for a single scenario (examples given in parentheses). Notes *Solid lines* indicate relations which always apply while *dotted lines* indicate relations that may or may not apply, as determined on a case by case basis

With these definitions in hand we can now provide a schematic representation of the overall integration of economic analyses within ecosystem service assessments for any single given scenario, as illustrated in Fig. 1.

Figure 1 provides a route map for our subsequent discussions. Starting from the top, we see the initial flow from primary and intermediate through to final ecosystem services. In many, but not all,[16] cases the latter services will then be augmented by human and manufactured capital inputs before generating welfare bearing goods. We then isolate the contribution of ecosystem services to the production of those goods as failing to do so negates the contribution of human and manufactured capital and so risks overstating the value of ecosystem services

[16] Hence the use of dashed lines to indicate relations that may not necessarily apply.

and undermining the credibility of such analyses. Once isolated, economic analyses seek to assess this value in monetary terms. While complex, this task permits the decision-maker to compare the value of ecosystem service benefits on an equal footing with the other goods (e.g. healthcare, education, etc.) which determine social wellbeing. However, as acknowledged in Fig. 1, at present not all of the benefits derived from ecosystem services are necessarily amenable to monetary valuation (e.g. environmentally related social norms; the spiritual value of the natural world; etc.). However, the drive of the UK NEA and the focus of this paper is to apply economic assessment as widely as possible and utilise alternative approaches[17] only as necessary, feeding these appraisals into the final assessment of wellbeing as constraints upon, or consequences arising from, alternative decision strategies. For any given scenario the final stage of the analysis is to assess the sustainability of a situation in terms of the stock of underlying ecological assets. As mentioned, we currently assume this away and return to this vital issue in the latter part of this paper. Altering the trend, environmental or policy drivers which define a scenario will result in changes to ecosystem service and manufactured or human capital inputs which will in turn feed through to different outputs of goods and hence altered values.

3 From Ecosystem Services to Welfare Bearing Goods: Scenario Analyses

Figure 1 illustrates the distinction between supporting primary processes and the final ecosystem services which contribute directly to the production of goods. However, an assessment needs to address both primary processes and final ecosystem services, as the status of the former provide vital information for evaluating the sustainability of systems (currently assumed to be sustainable), while the latter is required for analysis of the economic value of service flows. Clearly accurate quantification of natural assets and services is a pre-requisite for any credible ecosystem assessment. In the case of the UK NEA, a team of over 160 eminent natural scientists were assembled to quantify the status of ecosystem processes, assets and final ecosystem services they generate across the UK.[18] The natural science team was divided into two major groupings; those working to assess individual habitats (eight groups as follows: Mountains, moors and heathlands; Semi-natural grasslands; Enclosed farmland including arable and improved grasslands; Woodlands; Freshwater, wetlands and flood plains; Urban; Coastal margins; and the Marine environment) and those collating information across the former groups to asses trends and interactions in services (four groups as follows: Supporting services including primary and intermediate services; Provisioning services such as food, wood, etc.; Regulating services such as pollution assimilation, climate regulation, natural disease control, etc.; and Cultural services including those items to be quantified in non-monetary terms such as spiritual values, ecological education, etc.). In addition to these central groups a further team was assembled to ensure that the contribution of biodiversity

[17] Some argue that such issues require alternatives to the individual centred approach of economics, favouring instead group approaches to assessment (Wilson and Howarth 2002; Spash 2008). More compatible with the approach of economic analysis is the recent rise in research into direct measures of subjective wellbeing and happiness (Kahneman et al. 1997; Layard 2005; Oswald and Wu 2010). Here researchers model individuals' self-rated happiness as a function of their socio-economic circumstances and prevailing environmental conditions (Welsch and Kuhling 2008), thus permitting estimation of an implicit trade-off between the two.

[18] A substantial complication concerns international trade in ecosystem services (both imports and exports of ecosystem services and waste). While the UK NEA will conduct preliminary assessment of this trade the major focus is on domestically produced ecosystem services.

Table 1 Final ecosystem services and corresponding goods: Examples from the UK NEA

Final ecosystem service[a]	Principal related goods
Production of crops, plants, livestock, fish, etc. (wild and domesticated)[b]	Food, fibre, energy, genetic resources, industrial inputs, fertiliser, avoidance of climate stress, recreation and tourism, physical and mental health, ecological knowledge, etc.
Production of trees, standing vegetation and peat[b]	Timber, avoidance of climate stress, energy, noise regulation, recreation and tourism, etc.
Production of wild species diversity including microbes[b,c]	Natural medicine, disease and pest control, genetic resources, wild food, bioprospecting, recreation and tourism, physical health, ecological knowledge, etc.
Production of water quantity[b,c]	Potable water, Industrial use of water, flood protection, energy, recreation and tourism, physical health, ecological knowledge, etc.
Regulation of the climate[c]	Avoidance of climate stress, physical and mental health, ecological knowledge, etc.
Regulation of hazards; related vegetation and other habitats[c]	Coastal protection, erosion protection, flood protection, avoidance of climate stress, physical and mental health, ecological knowledge, etc.
Breakdown and detoxification of waste[c]	Pollution control, waste removal, waste degradation, physical and mental health, ecological knowledge, etc.
Purification processes[c]	Clean air, clean water, clean soils, physical health, ecological knowledge, etc.
Generation and maintenance of meaningful places; socially valued landscapes and waterscapes[d]	Recreation and tourism, physical and mental health, ecological knowledge, etc.

[a] As noted previously, other inputs (e.g. manufactured capital) may in some occasions be required to combine with final ecosystem services in the production of goods. Relating the final ecosystem services to the MA (2005) nomenclature:
[b] 'Provisioning' services;
[c] 'Regulating' services;
[d] 'Cultural' services. 'Supporting' services relate to primary ecological services

to all groups of services was assessed.[19] Further groups were added to allow analyses to be aggregated to specific spatial decision making levels, notably for each of the individual countries subsumed within the UK.

The economics team was organised to complement the natural science structure of the UK NEA with subgroups assigned to work with each of the habitat, service and biodiversity teams. An early task involving both the economists and natural scientists was to define the list of final ecosystem services and corresponding goods around which the economic analysis would focus. This clarified the distinction between final ecosystem service (e.g. trees) and economic good (e.g. timber). Table 1 provides an overview of some of the final ecosystem services and corresponding goods being considered under the UK NEA.[20]

Mapping and quantifying the linkages between primary processes, intermediate and final ecosystems services through to welfare bearing goods is one of the most fundamentally

[19] The special focus accorded to biodiversity reflects a concern that, because of its highly interactive and co-dependent role across multiple habitat groups, it might be overlooked if only dealt with at the habitat level.

[20] This classification appears similar and is conceptually compatible to that given by Haines-Young et al. (2009) in their recent report to the European Environment Agency on this issue.

important undertaking for an ecosystem assessment. Analysis of historic trends then permits a sustainability analysis of development to date. However, policy making requires future orientated assessments conducted through analyses of scenarios. The crucial outset point for any scenario analysis is the baseline from which all alternative scenarios are to be judged. A number of options exist for defining such baselines, including the present day position, extension of current trends or a Bayesian approach. Whichever approach is preferred, an obvious baseline would use prior trends and existing best estimates to construct the likely future path for ecosystem services taking into account three drivers: (i) available knowledge on environmental change (e.g. UK Climate Projections estimates of climate change), (ii) the likely path of market forces (e.g. OECD and FAO estimates of economic activity, commodity and fuel prices, etc.) and (iii) ongoing and planned policy initiatives (e.g. continued reform of the EU Common Agricultural Policy). The definition of alternative future scenarios has been a substantial focus of recent research (OST 1999, 2003; IPCC 2000; Environment Agency 2001; UKCIP 2001; Hulme et al. 2002; MA 2004; PMSU 2004; Pinnegar et al. 2006; Murphy et al. 2009; Foresight Foresight Land Use Futures Project 2010). However, like most ecosystem assessments, the UK NEA considers a number of alternatives reflecting multiple changes in the environmental, market and policy drivers outlined above. These reflect not only the best estimates underpinning the baseline, but variations due to uncertainty and a variety of policy response options. Scenarios can also be constructed to reflect aspirational objectives through which inspection of the changes required to achieve such ends can be undertaken.

4 Isolating the Ecosystem Service Contribution to the Production of Welfare Bearing Goods

As discussed in greater detail subsequently, the value of ecosystem services is generally assessed by looking at the welfare generated by the goods produced using those services (whether they generate use and/or non-use values). However, this does not imply that we can attribute all of the value of those goods to ecosystem services. To show this consider a situation in which a given ecosystem service input is lost. Outputs of associated goods may therefore reduce and in extreme cases fall to zero. However, even in the latter case we may well be in error if we assign all of the value of those goods to the ecosystem service. This is because those goods are often produced by combining ecosystem services with other resources such as manufactured or human capital. While the loss of ecosystem service inputs would almost inevitably lead to some net loss of value,[21] this value loss is not total as there will be some reallocation of these other resources to different ends. Assigning all of a good's value to its ecosystem service input assigns a zero value to all other inputs and ignores the potential for reallocation of the latter inputs to other productive ends. Such error would overestimate the value of ecosystem services in a way that risks undermining the credibility of assessments and a return to the default position of ignoring ecosystem values.

So, before we jump to consider the value of ecosystem services, we first need to consider the role they play in the production of quantities of goods. To do this we also need to consider and control for all the other inputs which affect that production. There are a number of ways to undertake such an analysis, but one which illustrates the underlying principles is to consider a 'production function' such as the following:[22]

[21] Both because of the loss of the ecosystem service itself and because one might expect that the other resources being used had been so allocated because this was the best way in which they could be used.

[22] For an introduction to production functions see Perman et al. (1996) and Chiang (1984).

$$q = q(m, n)$$

where q denotes the output of some good, m is a vector of manufactured and human capital inputs (e.g. labour) and n denotes a vector of natural capital inputs of which one is the ecosystem service we are focussing upon. There are a host of possible specifications of the production function (see Griffin et al. 1987) and the task of defining a given form typically involves considering both the empirical suitability of a form and how well it reflects reality. For the purposes of illustration we can consider the Cobb—Douglas (CD) production function (Chiang 1984) which has some properties likely to be reflected in ecosystem services. The CD function is as follows:

$$q = m^{\alpha} n^{\beta}$$

If there are multiple inputs involved in producing q (say 1 to M of the m inputs and 1 to N of the n inputs) then the CD production function becomes:

$$q = m_1^{\alpha_1} m_2^{\alpha_2}, \ldots, m_M^{\alpha_M} n_1^{\beta_1} n_2^{\beta_2}, \ldots, n_N^{\beta_N}$$

The CD function has a useful property when we take natural logarithms that:

$$\ln q = \alpha_1 \ln m_1 + \alpha_2 \ln m_2 + \cdots + \alpha_M \ln m_M + \beta_1 \ln n_1 + \beta_2 \ln n_2 + \cdots + \beta_N \ln n_N$$

where the coefficients $\alpha_1, \alpha_2, \ldots, \alpha_M, \beta_1, \beta_2, \ldots, \beta_N$ are known as the 'elasticities' of production of q of the respective inputs $m_1, m_2, \ldots, m_M, n_1, n_2, \ldots, n_N$. This means that, all else remaining equal, a one percent increase in input n_1 will lead to a β_1 percent increase in production of q. The value of these coefficients reflects available technology and will therefore change with technology but, broadly speaking, technical improvements will increase the value of the coefficients indicating that we obtain more output for a given unit of input.

The CD function (like most feasible production functions) exhibits the property that as we reduce the level of one input so we have to increase levels of one or more other inputs in order to maintain the level of output of q. The rate at which one input can be substituted for another while keeping output constant is known as the elasticity of technical substitution which we denote as σ. If $\sigma = 0$ then any reduction in an input will directly reduce output and cannot be offset by increases in other inputs.[23] This describes a 'strong sustainability' (Pearce et al. 1989) situation where the level of ecosystem service inputs from natural capital must be maintained to ensure long-term sustainability. At the other extreme, when $\sigma = \infty$ then the impact on q of an x percent decrease in one input can always be exactly offset by a y percent increase in another. Such as situation conforms to a 'weak sustainability' (*ibid.*) worldview in which there are no 'critical' inputs without which production is impossible and sustainability is ensured simply by ensuring that the sum of natural and manufactured capital is non-declining, rather than by needing to maintain any particular (e.g. natural) capital type.

Typically σ is neither zero nor infinite, lying instead between these extremes.[24] Here reductions in say natural capital inputs can be substituted for by increases in manufactured capital but in a manner such that, as the availability of n declines, so progressively greater quantities of m have to be employed to maintain output of q. Now where all inputs are priced

[23] An example is the Leontief production function where inputs can only be combined in fixed proportions such that an increase in just one input has no impact on production, while a decrease in one input cannot be offset by increasing another.

[24] The CD function provides such an example for which $\sigma = 1$. One of the less plausible aspects of the CD function is that it indicates that if and only if any input fell to zero then q would also cease. In reality such states might be unlikely to occur due to the very large levels of other inputs required to maintain q as any given input is run down toward zero.

at their true value then the progressive nature of this substitution relationship provides a break on the over-exploitation of any given resource. However, when some inputs (such as n) do not have market prices or those prices underestimate their true value then producers have an incentive to continually substitute n for m and drive down their costs irrespective of the loss of n (and associated true value) this will incur. This incentive persists even in areas of the production function where large inputs of n are needed to substitute for further reductions in m. Such incentives underpin the growing drive to value natural asset inputs as reflected in analyses such as the UK NEA.

Real world production relationships are often more complex than the CD case and may involve a plethora of inputs exhibiting a variety of output and substitution relationships within a single function (we consider an example subsequently). Despite this, for the reasons discussed previously, it is important to identify the contribution of ecosystem services to goods separately from those obtained from manufactured and other human capital. However, even when this is achieved we still face the tricky issue of valuing these ecosystem service inputs; an issue to which we now turn.

5 Valuing Ecosystem Services

Perhaps the source of most confusion in practical decision making concerning the environment are two commonplace terms which most people use interchangeably: 'value' and 'price'.[25] That they are not in fact equivalent is easy to demonstrate. Consider a walk in a local park. The market price of such recreation is likely to be zero as there are no entrance fees and anyone can simply walk in. However, the very fact that people do indeed spend their valuable time in parks shows that this is not a zero value good. It is clear to see that 'value' and 'price' are not necessarily the same thing. In fact price of some unit of consumption is simply that portion of its value which is realised within the market place. Now in some cases price may be a perfectly acceptable approximation to value, particularly where all the inputs to the production of a good are privately owned, that good is produced in a competitive market[26] and where there is not large scale intervention by governments or other authorities.[27] Indeed even when these latter distortions do arise economists can often adjust for their influence. However, as the park recreation example shows, market price can in some cases be a poor approximation of value, indeed this divergence can often be substantial and is a characteristic of many of the goods produced by the natural environment.

Economists have developed a variety of methods to for estimating the value of goods whose market prices are either imperfect reflections of that value or non-existent. These methods are discussed in detail through a variety of reviews and guidelines see, for example, Barbier (2007), Bateman (2007), Bateman et al. (2002a), Champ et al. (2003), Freeman (2003), Hanley and Barbier (2009), Heal et al. (2005), Kanninen (2006) and Pagiola et al. (2004). This is a substantial literature and we are unable to provide a detailed assessment within the confines of the present paper. Accordingly we present only a brief summary and critique of the methods overviewed in Table 2 with the reader referred to the above literature for an in-depth consideration. Furthermore we emphasise that, within this section we retain

[25] Another common and related confusion concerns the terms 'accountant' and 'economist'. Accountants are interested in market prices, whereas the true economist should be interested in values.

[26] Typically, the less competitive a market the more any individual producer can exerting pressure upon price.

[27] Interventions such as government subsidies or taxation can strongly distort prices away from their competitive market levels.

Table 2 Various valuation methods applied to ecosystem services

Valuation method	Value types	Overview of method	Common types of applications	Examples of ecosystem services valued	Example studies
Adjusted market prices	Use	Market prices adjusted for distortions such as taxes, subsidies and non-competitive practices.	Food, forest products, R&D benefits.	Crops, livestock, multi-purpose woodland, etc.	Bateman et al. (2003), Godoy et al. (1993)
Production function methods	Use	Estimation of production functions to isolate the effect of ecosystem services as inputs to the production process.	Environmental impacts on economic activities and livelihoods, including damage costs avoided, due to ecological regulatory and habitat functions	Maintenance of beneficial species; maintenance of arable land and agricultural productivity; support for aquaculture; prevention of damage from erosion and siltation; groundwater recharge; drainage and natural irrigation; storm protection; flood mitigation	Ellis and Fisher (1987), Barbier (2007).
Damage cost avoided	Use	Calculates the costs which are avoided by not allowing ecosystem services to degrade.	Storm damage; supplies of clean water; climate change.	Drainage and natural irrigation; storm protection; flood mitigation	Badola and Hussain (2005), Kim and Dixon (1986).
Averting behaviour	Use	Examination of expenditures to avoid damage	Environmental impacts on human health	Pollution control and detoxification	Rosado et al. (2000).

Table 2 continued

Valuation method	Value types	Overview of method	Common types of applications	Examples of ecosystem services valued	Example studies
Revealed preference methods	Use	Examine the expenditure made on ecosystem related goods (e.g. travel costs; property prices in low pollution areas).	Recreation; environmental impacts on residential property and human health.	Maintenance of beneficial species, productive ecosystems and biodiversity; storm protection; flood mitigation; air quality, peace and quiet, workplace risk.	See Bockstael and McConnell (2006) for the travel cost method and Day et al. (2007) for hedonic pricing.
Stated preference methods	Use and non-use	Uses surveys to ask individuals to make choices between different levels of environmental goods at different prices to reveal their willingness to pay for those goods	Recreation; environmental quality, impacts on human health, conservation benefits.	Water quality, species conservation, flood prevention, air quality, peace and quiet.	See Carson et al. (2003) for contingent valuation and Adamowicz et al. (1994) for discrete choice experiment approach.

Source: Adapted from de Groot et al. (2002), Heal et al. (2005), Barbier (2007), Bateman (2009a) and Kaval (2010)

our focus on just the present period but note that true values can only be fully captured within a multi-period assessment to which we turn subsequently.

Considering Table 2, as noted, market prices can in some cases provide an acceptable starting point for valuation.[28] However, adjustment should always be made to correct for market distortions such as taxes and subsidies (which are effectively merely transfers from one part of society to another) as well as for non-competitive practices (Freeman 1991; Nicholson et al. 2009; Dasgupta 2009). Related to this approach, is the factor input or production function method (see Barbier 2000, 2007; Freeman 2003 and Hanley and Barbier 2009). As discussed in the previous section, this examines the contribution of all of the inputs used to produce a good in terms of the value they add.[29] This approach can be applied to a range of market (consumption) goods but has also been used for valuing regulatory and 'protection' goods (examples of the latter including flooding and extreme weather protection).[30] All of these approaches infer values by examining linkages with (adjusted) market priced goods. This tactic is also used in the examination of potential value losses in terms of avoided damage costs or behaviour and expenditure intended to avert such damages.[31] However, we have excluded the use of restoration or replacement costs as a proxy for the value of ecosystem services. Although there are a few interesting examples of such studies,[32] many economists consider that such methods should be used with caution (Barbier 1994, 2007; Ellis and Fisher 1987; Freeman 2003; Heal 2000) due to the likelihood that restoration or replacement costs will bear little resemblance to the values they approximate.

The methods described above might appear straightforward. However, this is somewhat deceptive. Recall that the task of the economist is to estimate the value of goods in terms of the welfare they generate, rather than simply their market price. As mentioned it is only under a set of fairly restrictive assumptions that we can take market price as a direct estimate of value (recall the park recreation example) and the adjustment process from the former to the latter is far from straightforward. However, even this route becomes impassable for goods which are devoid of market prices such as outdoor, open-access recreation, or peace and quiet. Revealed preference methods provide an approach to the valuation of goods such as these where an individual can only enjoy some non-market environmental good through the consumption of some market priced private good. Here economists make use of the 'weak complementarity' concept introduced by Mäler (1974) to examine how much individuals are prepared to spend on the private good in order to enjoy the environmental good, thereby revealing the value of the latter. A number of variants of the revealed preference approach exist. For example, the travel cost method examines the expenditure and time that individuals are prepared to give up to visit environmental recreation areas. Similarly, the hedonic property price method examines the premium which people are prepared to pay in order to purchase houses in areas of higher environmental quality (e.g. quieter, less polluted neighbourhoods).

[28] The use of market prices within a green accounting approach to assessing ecosystem services is discussed by Cairns (2002).

[29] Examples of production function based valuations of ecosystem services include: multi-purpose woodlands (Barbier 2003; Boscolo and Vincent 2003; Nalle et al. 2004); marine nutrient balance (Gren et al. 1997; Smith 2007; Knowler et al. 2001; Knowler and Barbier 2005), pollination (Ricketts et al. 2004); power generation (Considine and Larson 2006); fisheries (Rodwell et al. 2002; Mardle et al. 2004; Sumaila 2002; Barbier 2003, 2007); watershed protection (Kaiser and Roumasset 2002; Hansen and Hellerstein 2007).

[30] Examples include the storm protection values of mangroves in Thailand (Barbier 2007) and hurricanes along the US Atlantic and Gulf coasts Costanza et al. (2008).

[31] Note that the averting behaviour method could also be viewed as a variant of the revealed preference approach discussed subsequently.

[32] See, for example, the study of the New York City drinking water source in the Catskills Mountains discussed by Chichilnisky and Heal (1998).

By controlling for other determinants (e.g. the number of bedrooms in a property) such purchases reveal the values people hold for these environmental goods.[33] However, while these methods have been widely applied, they have various drawbacks and limitations. They often require a number of assumptions to hold[34] as well as copious amounts of data and intensive statistical analysis.

While revealed preference techniques tend to be applicable to a relatively narrow range of goods, stated preference approaches such as contingent valuation and discrete choice experiment methods should, in theory, be applicable to a wide range of ecosystem service goods[35] and typically they are the only option available for estimating non-use values.[36] Such methods are defensible in cases where respondents have clear prior preferences for the goods in question or can discover economically consistent preferences within the course of the survey exercise. Where this is not the case then elicited values do not provide a sound basis for decision analysis. Such problems are most likely to occur for goods with which individuals have little experience and poor understanding (Bateman et al. 2008a,b, 2010a).[37] Therefore while stated preferences may provide sound valuations for high experience, use value goods, the further we move to consider indirect use and pure non-use values, the more likely we are to encounter problems.

While a number of solutions have been proposed for the problem of valuing low experience, non-use goods,[38] we have to consider those cases where such values cannot be established to any acceptable degree of validity. The question of what should be done in such cases has generated much debate, but one approach is the adoption of ecological standards (see subsequent discussion of 'safe minimum standards') to ensure the sustainability of resources (such as the continued existence of species) which are not amenable to valuation (Farmer and Randall

[33] Notice that the hedonic property price approach examines the value of a flow of services as capitalised within house prices. A related approach is to model the relationship between the price of land and its attributes. Examples of such 'Ricardian' analyses include Fezzi and Bateman (2010), Mendelsohn et al. (1994), Schenkler et al. (2005) and Seo et al. (2009).

[34] See, for example Randall (1994) on the travel cost method and Day et al. (2007) on the hedonic pricing method.

[35] The stated preference literature is vast but for a few examples focussed upon ecosystem services Naylor and Drew (1998), Banzhaf et al. (2006), Carlsson et al. (2003),Othman et al. (2004), Hanley et al. (2003), Hearne and Salinas (2002), Huybers and Bennett (2003), Mansfield et al. (2008), Naidoo and Wiktor (2005), Rolfe et al. (2000).

[36] Notice that we deliberately eschew the term 'intrinsic value'. The word 'intrinsic' is defined by the Merriam-Webster dictionary as "belonging to the essential nature or constitution of a thing". Therefore the intrinsic value of say an endangered British bird such as the bittern (Eaton et al. 2009) belongs to the bittern and cannot be accurately defined by another entity such as a human. Of course humans can and do hold values for bitterns. These can include the use value held by bird-watchers and the non-use values which a wider group hold for the continued existence of the bittern as a species. However, these are anthropocentric rather than intrinsic values. To claim that we have any knowledge of intrinsic values lends a confusing and erroneous air of moral justification to assessments. Although some would argue for notions of human assigned intrinsic values (e.g. Hargrove 1992), from an economic perspective much so called 'intrinsic' values would instead be reclassified as non-use existence values. Arguably true intrinsic values (e.g. the value of the bittern to the bittern) could be protected by a property rights approach which makes it illegal to harm the species concerned. However, such rules are likely only to be enacted when they are actually supported by anthropocentric non-use values.

[37] A related problem is where variants of the stated preference approach provide survey respondents with heuristic cues regarding response strategies (Bateman et al. 2009a).

[38] One approach is to use more naturalistic approaches to the description of low experience goods; for example using virtual reality software to convey images of landscape goods (Bateman et al. 2009b). Others have proposed the use of intensive valuation workshops (Christie et al. 2006). However, such techniques are prone to reliance upon small, unrepresentative samples which, after such intensive experiences cannot be taken as reflecting general preferences.

1998). This would not negate the need for economic analysis which would still play an important role in the identification of cost-effective approaches to ensuring the maintenance or the provision ecologically sustainable levels.[39] A related strategy, the implementation of offsetting compensatory projects validated for their ecological suitability (Federal Register 1995), would also generally require such cost-effectiveness analysis.[40]

Despite the various caveats outlined above, probably the most serious problems facing the effective and robust valuation of ecosystem services are gaps in our understanding of the underpinning science relating those services[41] to the production of goods and the paucity of valuation studies and available data regarding the values of these goods. While interdisciplinary efforts between natural scientists and economists are crucial to addressing the former problem, the lack of valuation studies has lead to the ongoing development of methods to transfer values from existing 'source' studies to desired 'policy' applications.

The value transfer literature[42] embraces a number of approaches. The simplest technique is to search for a prior source valuation study which addresses a good and context which approximates that of the policy application and apply the value from the former to the latter.[43] This simple approach, often referred to as mean value transfer (because typically it is the average value which is transferred) is defensible provided that source and policy good and context are highly similar. However, the limitations of source valuation studies mean that this often not be the case. In such cases one option is to attempt to adjust the source values by incorporating differences between the source and policy contexts (e.g. differences in good characteristics, changes in valuing populations and their characteristics, different use costs or substitute/complement availability, etc.). Although some studies have attempted to affect such adjustments through expert assessment, a quantified adjustment process is clearly preferable if executed carefully.[44] One variant of the latter approach is to conduct a meta-analysis of results from previous studies, relating values to the characteristics of those studies and the goods and contexts valued. Such an analysis typically yields a regression model linking values to the characteristics captured in the available source data. The analyst can then apply the characteristics of some policy case to this model to estimate the relevant value. So, for example, Brander et al. (2006) conduct a meta-analysis of over 190 wetland valuation studies, providing 215 value observations. Significant determinants of value include, amongst

[39] For an example of a cost-effective approach to species preservation see Bateman et al. (2009c) and contrast this to the highly variable stated preference values for these projects given in Bateman et al. (2010a).

[40] An exception being cases where there is only one option available for delivering sustainability, in which case that option becomes a fixed constraint on any development policy (Barbier et al. 1990).

[41] Two problems are particularly highlighted: (i) the availability of quantified data on changes in the provision of services over time and space under different scenarios; (ii) quantified understanding of the interactions between ecosystems and their services, particularly under novel general stressors such as global climate change. These issues will require concerted action and high degrees of collaboration between the natural and social sciences.

[42] Examples of value transfers (sometimes called benefit transfers although this is confusing as these techniques can also be applied to costs) and related meta-analyses for environmental goods include Barton 2002; Bateman et al. 2009d, 2010b; Bateman and Jones 2003; Bergland et al. 1995; Brouwer 2000; Brouwer and Spanninks 1999; Brouwer et al. 1999; Carson et al. 1996; Desvousges et al. 1992; Downing and Ozuna 1996; Johnstone and Duke 2009; Johnston et al. 2005, 2006; Leon-Gonzalez and Scarpa 2008; Lindhjem and Navrud 2008; Moeltner et al. 2007; Muthke and Holm-Mueller 2004; Navrud and Ready 2007; Ready et al. 2004; TEEB 2009; Zandersen et al. 2007.

[43] Transfer databases such as The Environmental Valuation Reference Inventory (EVRI, available at www.evri.ca) have been developed to assist the search process for such applications.

[44] This caveat is important. Poor or incomplete quantitative adjustments can generate greater errors than simple mean value transfers (see demonstration by Brouwer and Bateman 2005). However, accurate, quantified and full adjustments outperform simple mean transfers (Bateman et al. 2009d).

others, wetland size, the ecosystem services and goods provided and population characteristics related to size and income.[45] By quantifying these variables for some policy application, analysts can use the parameters of the meta-analysis model to estimate values for the policy case.

An alternative and potentially more flexible and sophisticated approach to adjusting from source to policy values is to estimate and transfer a spatially explicit value function. This approach can be applied to the estimation not only of unit values but also to predicting the quantity change to which they are applied. For example, Bateman et al. (2010c) collect data on the location and frequency of trips to river sites across a large area of north-eastern England (chosen because of its diversity of river qualities and locations, and the variability of population density and socioeconomic characteristics). By interviewing a large sample of households and applying a variant of the travel cost model[46] the analysts estimate a model which predicts how the number and value of visits responds to a variety of spatially variable factors. These include both the location and quality of rivers, but also of substitute and complementary attractions and the density and socioeconomic characteristics of the population. Once such data is gathered its analysis yields estimates of the parameters describing the relationship between each of the above factors and the number and value of visits. Transferral of such functions is then relatively straightforward as secondary data sources can provide information on all of the predictors in the model for any area in the country (e.g. all locations and distances are provided by analysis of map data while official environment agencies hold data on river quality and census bodies store population characteristics).[47] Extracting such data for areas of policy interest and applying these to the estimated parameters yields the desired value and demand estimates.[48]

Whatever method is chosen for valuing preferences for an environmental non-market good, it must be capable of capturing the complex contextual aspects of such values. One of the most complex yet important aspects of such values is that even when, as we continue to assume, overall stocks are at or above sustainable levels the size of that stock may affect the value of changes in flows. This can be illustrated in part through reference to the highly cited study by Costanza et al. (1997) which attempted to provide value estimates for the total stock of all ecosystem services globally. While their paper very substantially raised awareness of the application of economics to ecosystem assessments, particularly within the natural science community, the focus upon valuing total stocks has been criticised on a number of grounds (e.g. Heal et al. 2005).[49] In particular, very few policy decisions relate to total losses of ecosystem services. Instead policy changes require an understanding of the value of changing a single unit of a stock. Economists refer to this as the 'marginal' value of the ecosystem service in question. Of course if the value of a marginal unit is constant then it is straightforward to go from valuing a single unit to valuing whatever number of units

[45] A complicating factor in such meta-analyses is the significant effect exerted upon values by the choice of valuation methodology in the source studies (see also Bateman and Jones 2003).

[46] Specifically the travel cost random utility model (RUM) (see Bockstael and McConnell 2006; Champ et al. 2003; Haab and McConnell 2002).

[47] As discussed in Bateman et al. (2002b, 2006) and Troy and Wilson (2006), a geographical information system (GIS) greatly eases the computational tasks underpinning related operations such as distance calculations and can readily be adapted to undertake the value transfer exercise.

[48] While the development of these various value transfer methods has allowed researchers to increase the usefulness of the existing literature (TEEB 2009; Bateman et al. 2010b), there remains a significant information gap with respect to source valuations which will only be addressed through a concerted and efficiently designed programme of research (Carpenter et al. 2009; Nicholson et al. 2009).

[49] Note that while they do not provide solutions to these problems, Costanza et al. (1997) are well aware of these issues and raise these within the discussion of their findings.

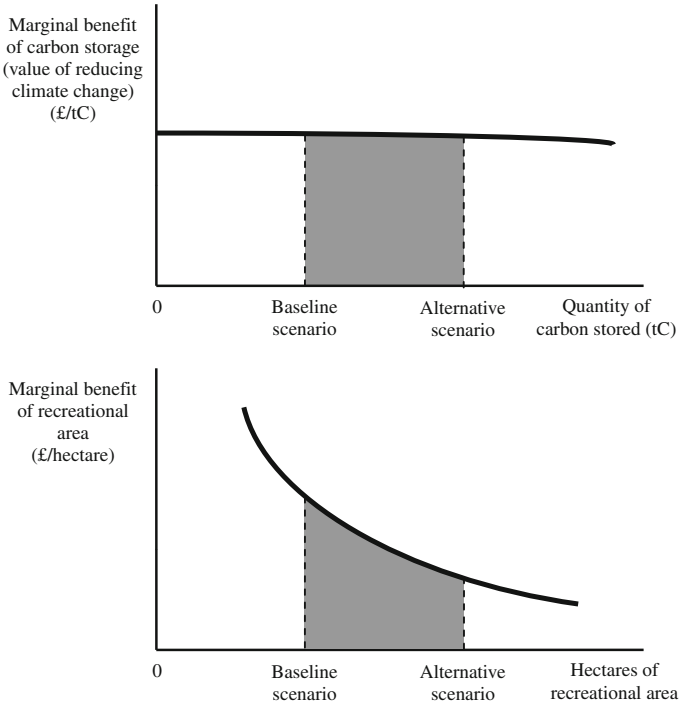


Fig. 2 Marginal benefit curves for two goods

a given policy will create or destroy. However, an interesting phenomena is that for many goods and services, marginal values will change with the total size of the stock, even when the overall stock level is above sustainable levels (as presently assumed).

Figure 2 illustrates the relevant point here by contrasting the two cases: the first concerning the marginal benefit (i.e. the per unit value) of reducing climate change by increasing carbon storage; the second showing the marginal benefit of increasing the area of recreational greenspace. In both cases we postulate a situation where there is a policy which changes land use so as to increase the provision of both carbon storage and land for recreation (e.g. through the creation of wetland areas suitable for peat accumulation and hence carbon storage as well as recreational visits).

The upper panel of Fig. 2 shows a (virtually) constant level for the marginal benefit of carbon storage throughout the range of feasible projects confined to some area of the UK. This reflects the simple fact that, using existing technologies whereby the bulk of carbon storage is held in living biomass and soils, the UK is simply not big enough to capture sufficient carbon to significantly reduce the problem of climate change to the level where the marginal benefits of further carbon capture change. Only if carbon sequestration were to be undertaken on a truly global scale would it begin to significantly affect the potential for damaging climate change and hence reduce the marginal value of further carbon capture. Here then, the total benefit value of the envisioned provision change is estimated by multiplying the (constant) marginal benefit of carbon capture by the increase in provision between the baseline and alternative scenario.

A more complex situation is shown in the lower panel which concerns increases in the area of recreational land. While initial provision of such area may be highly valued, once

that is provided, further (marginal) units of land generate progressively lower increases in recreational value. The Brander et al. (2006) meta-analysis of valuation studies mentioned previously provides a good example of such a case with per hectare values diminishing as the overall size of a wetland area increases. This pattern of diminishing marginal values is a characteristic of many goods (even carbon capture would exhibit such a pattern once climate change began to be significantly ameliorated).[50]

The lower panel of Fig. 2 also illustrates why it may be unwise to attempt to estimate the total value of ecosystem stocks rather than the value of specified changes. A total value would be given by integrating the marginal value curve back to a level of zero provision. However, such a situation (e.g. the disappearance of all recreational land) may be highly unlikely to occur. Equally importantly it moves the calculation through areas of the marginal benefits curve which are entirely unsupported by data. Extrapolation out of the range of existing data is likely to generate unrealistically high values. One common alternative to this approach is to use the current level of marginal benefits and hold this constant for the calculation of total values. However, just as the former approach is likely to generate over-estimates of value, this latter method ignores the shape of the marginal benefits curve and is liable to lead to underestimates of total value. Both options are unattractive and unnecessary. The focus upon changes in value between feasible scenarios is much more relevant for decision purposes. Accordingly this is the approach adopted for the UK NEA which argues that, for the valuation of any good we require:

(i) Understanding of the change in provision of the good under consideration (i.e. the change in the number of units being provided) given changes in the environment, policies and societal trends;
(ii) A robust and reliable estimate of the marginal (i.e. per unit) benefit value; and
(iii) Knowledge of how (ii) might alter as (i) changes.

The major challenge facing these requirements is the present state of knowledge regarding all of these issues. In the absence of full information (which will often be the case), both ecosystem assessments and accompanying economic analyses need to explicitly acknowledge uncertainties in their work and convey these to decision makers in clear terms. This in itself poses a substantial challenge which goes beyond the remit of this present paper.

6 Illustrations: Land Use Change

In this section we illustrate the various principles laid out previously, highlighting the issues of identifying the contribution of certain final ecosystem services to the production of both market and (subsequently) non-market goods and their economic valuation. We also use this example to illustrate the importance of incorporating the spatial complexity of the environment into analyses,[51] including the multiple values that can be generated by a given change and the conflict of values generated by mutually exclusive goods.

[50] The two panels of Fig. 2 also reflect the role of location in determining values. While the benefits of storing a tonne of carbon are spatially unconstrained (all individuals gain from this good), the benefits of increasing the size of a given recreational area are highly spatially confined, being disproportionately captured by those who live near to the site. This of course means that the spatial location of recreational sites near to population centres can substantially increase their value. Bateman et al. (2006) discusses the concept of 'distance decay' in such values. Note also that this raises the possibility of localised losses of stocks occurring even when regional, national or global stocks are maintained. This is likely to generate high spatial specificity in marginal values.

[51] For further discussion of both spatial and temporal complexity see Rodriguez et al. (2006).

Table 3 Land use share equations parameter estimates for cereals

Variable	Estimated coefficient
$\text{Price}_{\text{cereals}}$	0.134 ∗ ∗∗
$\text{Price}_{\text{fertilizer}}$	−0.111 ∗ ∗∗
Set aside rate	−0.425∗∗∗∗
Environmentally sensitive area (ESA) share of grid square	−0.033∗∗∗∗
Park share of grid square	−0.019 ∗ ∗∗
Urban share of grid square	−0.028 ∗ ∗
Share of agricultural land with slope higher than 6°	−0.087 ∗ ∗∗
Coastal (dummy)	−0.357
Mean elevation	14.170∗∗∗∗
Mean elevation squared	6.333 ∗ ∗∗
Autumn machinery working days	4.174∗∗∗∗
Autumn machinery working days squared	−1.283 ∗ ∗∗
Mean potential evapotranspiration	6.727 ∗ ∗∗
Mean potential evapotranspiration squared	−2.773 ∗ ∗
Median duration of field capacity	−4.794∗
Median duration of field capacity squared	16.670 ∗ ∗∗
Total number of degree days in the growing season	−4.228 ∗ ∗∗
Total number of degree days in the growing season squared	2.571 ∗ ∗
Average annual rainfall	−3.726
Average annual rainfall squared	−1.269
Trend	0.015
Constant	38.040∗∗∗∗

Notes: Due to the limitations of space the equations relative the other land uses and the parameters corresponding to the interactions of the environmental factors are not reported in the Table, but are available on Fezzi and Bateman (2009). * t-stat > 2; ** t-stat > 3; *** t-stat > 4; **** t-stat > 10

In order to provide a common empirical theme to this illustration we consider the wide-ranging issue of land use change as this combines all of the topics covered so far in this paper. Considering first farmed land, we apply the structural model of agricultural land use developed by Fezzi and Bateman (2009). This model draws upon highly disaggregated 2 km resolution data for the entire area of England and Wales over a period spanning more than 40 years. The model shows how profit maximising land use alters in response to changes in the physical environment (e.g. climate change), market conditions (e.g. prices and costs) and policy (e.g. revisions of the EU Common Agricultural Policy). In estimating this model, rather than adopt a relatively inflexible function form such as a Cobb-Douglas a more sophisticated specification[52] was employed to yield land use share estimates for all major agricultural arable and pastoral activities. Table 3 reports just one of these; a model of the share of each 2 km square of England and Wales which is devoted to cereals.

The model reported in Table 3 is, in effect, a spatially explicit transfer function, allowing us to predict the amount of land devoted to cereals in any location for which we have the

[52] A normalized quadratic function was adopted as this has been widely applied in agricultural economics for modelling joint (in input) multi-output production processes (Moore et al. 1994; Oude and Peerlings 1996; Guyomard et al. 1996; Arnade and Kelch 2007).

predictors shown (which, for the UK, are held for all areas of the country). The model also allows us to isolate the contributions of a variety of ecosystem services to the production of the good in question. This is shown by the estimated parameters corresponding to each of the ecosystem services contributing to the decision to produce cereals. Furthermore we can use these parameters to investigate the impacts of a change in any or all of these ecosystems services and/or the other determinants. This change might be derived from some external model forecasting future trends or from an explicit scenario analysis reflecting expert estimates and/or policy aspirations.

We can illustrate the use of such models by varying just one of these ecosystem services; the climate. We might obtain predictions of future climate from some existing model (e.g. the IPCC model summarised in Pachauri and Reisinger 2007). However, for purely illustrative purposes we can adopt a simple climate change scenario, holding constant all other land use determinants (prices, policy, urbanization, other ecosystem services, etc.) and increasing daily average temperature by 1°C; an increase which is expected to occur in the UK by about 2030 (Murphy et al. 2009).[53] This in turn alters the total number of degree days in the growing season captured in our model. Applied across all of the major agricultural land uses embraced in the model we find a highly heterogeneous pattern of response across different areas of England and Wales. This is hardly surprising given that, despite the relatively modest size of this area it embraces a great diversity of farming conditions and agricultural activities. The resultant spatial pattern of changes in the production of cereals is illustrated in the left hand panel of Fig. 3. This shows the area of cereals increasing in the northern parts of the country, where the warmer temperature will be beneficial to yield, and decreasing in the south where it is substituted for by other activities.

While climate change induced shifts in land use pattern are clearly important (informing us about impacts upon food production, landscape changes and a host of related impacts), from the perspective of economic analysis we need to monetise these changes. The right hand panel of Fig. 3 takes into account all of the shifts in land use driven by this simple climate change (i.e. including the substitution across all of the different agricultural activities) and provides a simple market price assessment of the financial impact[54] of this scenario.

The analyses of Fig. 3 are useful in that they demonstrate the principles of isolating the contribution of a given ecosystem service (here climate) within the production of goods. They also amply demonstrate the importance of incorporating spatial complexity within such analyses; as can be seen the impacts of climate change vary substantially (and from negative to positive) across England and Wales. However, this analysis does not go far enough in that it needs to move from market price to assess its true social value as well as considering the wider impacts induced by land use change. In order to illustrate this we present an analysis of a part of the above area, focussing upon the country of Wales, and draw on the work of Bateman et al. (2002a, 2003) and Bateman (2009b).

Figure 4 presents, in the left hand panel, the market price of Welsh agricultural output[55] and, in the right hand panel, its economic value. Both have been estimated by transferring a spatially explicit model as discussed previously. Results from this exercise show the

[53] The simple scenario used in the present paper is purely for illustrative purposes and we ignore issues such as monthly variation in the rate of climate change, precipitation impacts and linked environmental changes (e.g. in evapotranspiration rate, etc.). Other work by the authors relaxes these assumptions and addresses the projections given in work undertaken by the UK Climate Projections project (Murphy et al. 2009).

[54] This is here assessed using the commonly adopted measure of Farm Gross Margin (FGM) which is defined as the difference between revenues from agricultural activities and associated variable costs (FBU 2009; Defra 2009). The present illustration uses FGM figures for 2004.

[55] The figures actually present results for the dairy sector. Other sectors are presented in Bateman et al. (2003).

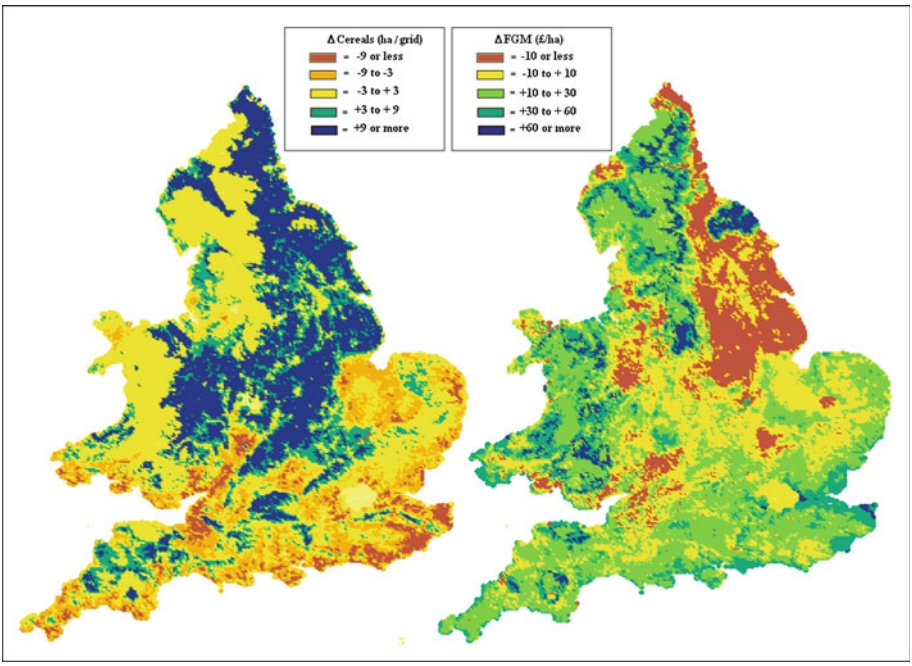


Fig. 3 Change in cereals (*left hand panel*) and market price assessment (*right hand panel*) under a simple climate change scenario

importance of spatial factors in determining values, with these being greatest in lowland coastal and eastern areas and lowest along the mountain chain which runs from north to south through central Wales. As can be seen, in this case, true values are below market prices because the latter are inflated by a series of subsidies and market interventions. This illustrates an important point that while economic values are often above market prices (due to the latter ignoring the elements of value that bypass the market), this relationship can be reversed, most noticeably when (as here) governments intervene to distort markets and raise the price of goods so as to increase their production.

While the valuation exercise illustrated in Fig. 4 is helpful, it is insufficient to guide land use decision making, most noticeably because it ignores 'opportunity costs'; the value which could be generated by alternative use of this land. Within the predominantly rural area of Wales one obvious alternative use is as multipurpose woodland. This would generate both market priced outputs, such as timber, and a variety of non-market values, of which we will just consider carbon storage and open-access recreation as illustrations.[56]

In assessing the quantitative change in provision which an alteration in land use from agriculture to multi-purpose woodland would induce, we employ spatially explicit transferable production functions similar to those discussed with respect to agriculture in our prior cereals example. In the case of timber such value functions take into account a host of spatially explicit data such as soil type, temperature, rainfall, topographic shelter, aspect, etc.

[56] Conversion from agriculture to woodland increases carbon storage in living biomass and in virtually all soils except peatlands where tree planting causes peats to dry out and emit carbon; similarly in the UK most state forestry is open-access and hence has higher recreational value than enclosed farmland. Other social benefits of afforestation not considered in this illustration include water quantity and quality regulation and the provision of biodiversity habitat.

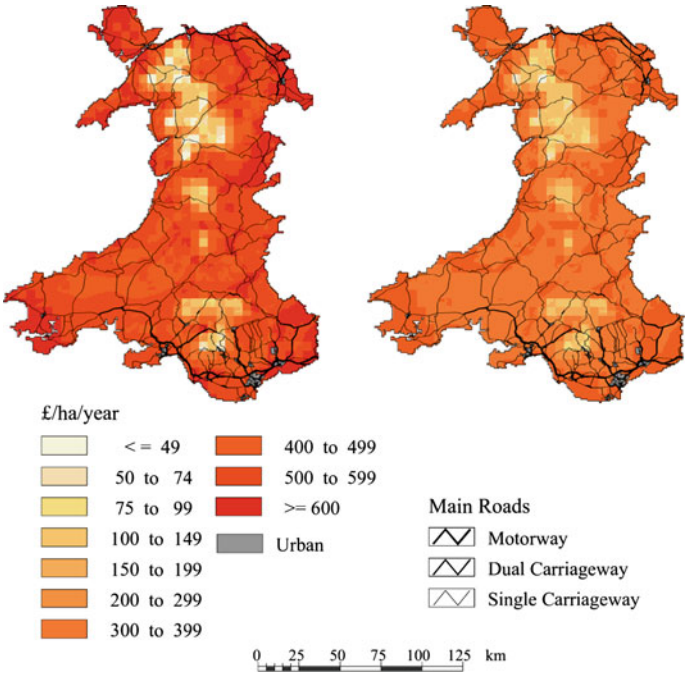


Fig. 4 Agricultural production expressed via its market price (*left hand panel*) and economic value (*right hand panel*), the latter adjusting for subsidies, market distortions, etc. *Source* Adapted from Bateman et al. (2003)

all of which can be derived from existing, high resolution national coverage databases. This allows us to estimate tree growth and timber production for different species and management regimes under baseline or alternative scenarios. Net carbon storage can then be modelled with respect to the growth of biomass, the storage or emission of carbon from different soils and the emission of carbon from felling and products (see Bateman et al. 2003 for details).

In attempting to value items such as timber production and carbon storage (and indeed all others being considered here) we now have to relax our focus upon a single period as it is obvious that these processes can take many years to develop yielding a flow of benefits (and costs) occurring at different points in time.[57] Bringing together both the present and future allows us to expand and complete the economic concept of value. Returning to the banking analogy referred to early in this paper, the economic perspective argues that the value of some stock of wealth should be measured by assessing the present day worth of all the future flows of income which that wealth is expected to generate. Economic analysis achieves this by the process of discounting; a procedure which reflects the general preference for benefits which occur in the shorter rather than longer term. We consider the issue of discounting in a little more detail subsequently but the important issue for the current illustration is that it converts a series of future benefit flows into a single value measured in present day terms.

[57] For detailed description of these analyses see Bateman et al. (2003). The timber valuation analysis takes into account subsidies and other market distortions. Both this and the carbon valuation take into account forecast changes in the real value of these goods (discussed in the sustainability section of this paper), while the carbon analysis considers carbon storage in above ground, species-specific biomass and its post-felling emission profile (adjusted for species and end-uses) as well as carbon flux within soils as a result of land use change.

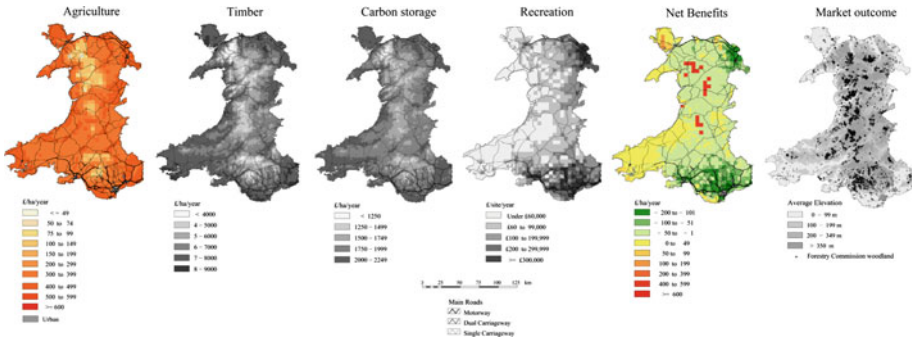


Fig. 5 *Shadow values* for various land use benefits. *Sources* Adapted from Bateman et al. (2002b, 2003) and Bateman (2009b). Values shown as annual equivalents, the discounted sum of which is the *shadow value*

This is known as the 'shadow value' of these income flows; a term which applies just as well to the benefits arising from ecosystem services.[58]

So, if we were undertaking an economic analysis of the timber benefits of a woodland we might assess its stock size in hectares and its flows of timber in cubic metres. However, to assess its shadow value (in pounds) we consider the present value of those future flows of timber. Of course we would typically want to consider all benefits, not just timber, and would apply the shadow value approach to each of these benefits.[59] Figure 5 attempts to illustrate such an analysis through a series of maps, all but the last of which details the shadow value of the various benefit streams which arise from the land use decision under consideration.[60] The first map relates to current agricultural land use. The next three maps (given in grey-scale to highlight their common origin) illustrate respectively the timber, carbon storage and recreation values which would arise from changing land use into multipurpose woodland.[61] The fifth map details the shadow value of the net benefits arising from subtracting the sum of woodland values from the agriculture value. Here the green areas indicate locations where woodland provides a higher shadow value than agriculture while yellow and red areas indicate locations where agriculture provides the highest values. It is interesting to note that the areas which generate the highest shadow values from conversion into woodland are in the north-east and south-east, a result which reflects the high populations in these areas and consequent elevated recreation values arising from afforestation. In contrast the most negative shadow values from such conversion are shown by the red areas corresponding to upland peats where afforestation causes major losses of soil carbon. The geographic distribution of net benefit

[58] I am particularly indebted to Daan van Soest for comments regarding this issue.

[59] Subsequently we discuss how values may change when stocks approach unsustainable levels. However, even above those levels some individuals may have preferences and hence values regarding a few stocks such as the existence values associated with populations of some species.

[60] For ease of comparison each value has actually been calculated as an annual equivalent, the discounted sum of which is the shadow value. Therefore the shadow agricultural values given as the first map in Fig. 5 are virtually identical to those shown in the right hand panel of Fig. 4. For details see Bateman et al. (2003).

[61] Each of these maps shows the change in value which this benefit stream generates. For the timber and recreation case this is a change from a zero prior value. However, for the carbon storage case this derives from the difference in storage generated by woodland as opposed to agriculture. The recreation analysis shown uses a transferable trip generation function to estimate the number of visits which would occur if a forest was created in a given location and a separate meta-analysis to estimate per-visit values (with aggregate values being given by multiplying per-visit values by the estimated number of visits). For a superior RUM approach to travel cost analysis see Herriges et al. (2004), Egan and Herriges (2006) and Bateman et al. (2010c).

shadow values is in sharp contrast with the actual distribution of forests as illustrated in the final map. The latter is driven by market forces alone and hence ignores the carbon storage and recreation values instead being driven solely by the market values of agriculture (the left hand panel of Fig. 4) and forestry (the timber values alone). Here agriculture dominates all lowland areas pushing forestry up the hill to low productivity areas where land prices are lower.[62] This results in a distribution of woodland which is in marked contrast to its true shadow value; a finding which underlines the importance of using shadow values rather than market prices to lead decision making.

7 Discounting

Towards the end of our illustrative land use study we relaxed our single period assumption in favour of considering the value of ecosystem services across multiple periods through the process of discounting. Here the shadow value of each good is assessed as the discounted present value of the stream of net benefits which are expected to be received into the future. For example, considering the shadow value of carbon storage we multiply the quantity of storage delivered in each year from the present onwards (ideally we should not be limiting appraisals by any arbitrary date) by an estimate of the social cost of carbon (see, for a discussion, Tol 2005).[63] These values are then discounted and summed back to the present day. The reason why this does not generate an infinite sum is that the discounting process progressively reduces the present day value of future costs and benefits with this reduction increasing in intensity the further into the future we go.

The discounting procedure is based upon both theoretical and empirical arguments that individuals have a preference for receiving benefits sooner rather than later. This means that shadow values encapsulate within them conceptions of the impact of changes in the stock of all assets (including natural assets) upon intergenerational wellbeing. However, both the form and rate of the discounting procedure are the subject of intense controversy. This is nowhere more evident than in the debate surrounding the recent Stern Review on the economics of climate change (Stern 2007). Subsequent argument has focused on the evidence that underpinned the central conclusion of the Review that the "the benefits of strong, early action far outweigh the economic costs of not acting" (p. vx). In particular, the focus of much of this discussion has been on the way in which this conclusion was driven by choices made in setting the social discount rate[64] including all of the fundamental reasons for discounting: pure time preference, the utility value of future increments in consumption and the extent to which it can be assumed that future consumption will be higher than consumption today (see, for example, Nordhaus 2007; Weitzman 2007; Dasgupta 2007).

A critical element of this debate centred on whether, in selecting the social discount rate, a descriptive approach or a prescriptive approach should be used (Dietz et al. 2007), a distinction which also can be found in IPCC (1996). Put another way, should investments in natural assets be appraised respectively in the light of information about preferences for the future as revealed in actual economic decisions or is there room for the practitioner to make explicit

[62] This includes pushing forestry onto peatlands resulting in some woodlands which are net contributors to global warming (Adger et al. 1992).

[63] Note that it is perfectly feasible that the 'real' value of a tonne of stored carbon need not be constant in all years. If the level of atmospheric CO_2 rises over time so the marginal damage of an additional unit of emissions is likely to rise and with it the value of avoiding that emission. 'Real' vales adjust for any intervening inflation. All cost-benefit analyses should be conducted using such real values.

[64] This is the rate relevant for decisions made on behalf of and reflecting the wishes of society. It differs and is typically markedly lower than the market discount rate which reflects private investment decisions.

moral judgments about intergenerational equity. Stern (2007) opts for the latter approach and, as a result, adopts a very low discount rate giving a relatively high weight to future costs and benefits. The substance of Nordhaus (2007) and Weitzman (2007) is that there is, on the face of it, very little evidence that this moral reasoning is reflected in people's actual behaviour and choices and, thus, the empirical evidence suggests that the pure rate of time preference should take a higher value. Resolving such debates is far from straightforward, however, and faces profound questions on which, to quote Beckerman and Hepburn (2007) "… reasonable minds may differ" (p. 198).

Interestingly, recent discussions surrounding discounting have also broken new ground with the growing recognition that some environmental problems such as climate change are truly 'non-marginal' in the sense that this problem could end up shifting the global development path, say with 'business as usual' emissions of greenhouse gases possibly leading to considerably lower future consumption levels than now (Weitzman 2007; Hoel and Sterner 2007). Indeed the corresponding notion that the socially appropriate discount rate for short term effects might differ from that relevant to long term impacts (such as climate change) has caught hold in practical guidance (see Treasury 2003 for example and for review, support and critique see Groom et al. (2005), Dietz and Hepburn (2010) and Dasgupta (2001) respectively).

8 Sustainability

Following our shift of focus from single to multi-period assessment, we can now relax our final assumption that underlying stocks of ecological assets are held at sustainable levels[65] (or equivalently that the rate of extraction or 'harvesting' of ecosystem service flows is sustainable).[66] This move from a sole focus on flows to also considering stock is analogous to moving from only thinking about expenditure to also considering the wealth which supports it. Accordingly the economic literature surrounding this issue develops the notion of "Comprehensive Wealth" which considers the ecological stocks from which all ecosystem service flows are generated (Dasgupta and Mäler 2000; Arrow et al. 2007; Mäler et al. 2008; Dasgupta 2009). Clearly any economic analysis of ecosystem services has to appraise the impact of potential stock depletions in order to assess the sustainability of given states. Most of the literature has focussed upon assessing historic development paths through adjustments of national income accounts (Bartelmus 2001, 2008; United Nations 2003; Hamilton and Ruta 2009), however, given the future orientated emphasis of the present paper we focus upon approaches appropriate for project appraisals of alternative options. Specifically we consider three potential strategies for incorporating sustainability concerns into such appraisals: (i) assessment of how future depletion of ecosystem stocks might increase the marginal shadow value of corresponding services; (ii) incorporation of the insurance value of maintaining

[65] It is important to remember throughout the following discussion that the economic notion of sustainability is rooted in values rather than simply physical stocks. This means that judgements regarding intergenerational wellbeing (including those reflected within the discounting process) are relevant here (see, for example, Hamilton and Clemens 1999; Dasgupta and Mäler 2000; Hamilton and Withagen 2007; Mäler et al. 2009).

[66] Recall our earlier note that it is not the physical constitution of these assets (e.g. coal stocks) which is the issue but rather the sustainability of the services they provide (e.g. energy). Some resources are inherently non-renewable and therefore all usage is depletive in nature. However, the proceeds of such use can be invested so as to maintain service levels (e.g. using oil-fuelled economic growth to develop renewable energy alternatives). A significant issue is that the costs associated with using different forms of service may differ substantially. For example extraction costs for crude oil are far lower than those from oil shale (Strahan 2009). This may in turn have implications for intergenerational wellbeing.

ecosystem resilience and; (iii) the use of safe minimum standards as a means of preserving stocks of ecosystem assets.

8.1 Increasing Future Marginal Shadow Values

Discounting is the process of converting benefits and costs occurring at some future date into their present day value. We remove the impact of inflation from this procedure and deal with the real (inflation adjusted) values arising at different points in time. However, from the seminal work of Krutilla and Fisher (1974) we know that there is no reason to suppose that these real values will stay the same throughout time. One factor which might affect the future marginal shadow value of ecosystem service flows is a change in the ecosystem asset stock from which that flow derives. Even then such stock reductions will only affect future values if the asset is reduced below some critical threshold which affects the sustainability of those flows. So, for example, consider the case of quartz, which is one of the most abundant minerals on earth. It is a major source of silicon which in turn is a vital input to numerous goods ranging from glass to food additives through to optical fibres and nanotechnology devices. However, due to the abundance of its stocks, any feasible pattern of depletion will not change the future value of silicon. But this is not the case for the ultra-pure mineral quartz used for computer processor manufacture for which stocks are only found in a very few locations (such as the famous mines at Spruce Pine, North Carolina). Here stocks are finite and continued exploitation could feasibly deplete these to a 'threshold' level below which the flow of such ecosystem services becomes reduced. One would expect that this would begin to increase the marginal value of those flows. While this example concerns goods which are traded in markets, the same principles should apply to non-market environmental goods except that now these effects will not be reflected in market prices but should be incorporated within the marginal shadow values used for economic analyses of ecosystem services.

The particular path that future marginal shadow values might follow depends on a number of factors and there are complicated issues to resolve about the empirical relationship between asset stocks, the flow of services and the way in which these services are valued at different stock levels (Pascal et al. 2009). However, a general point worth making is that if it is difficult to replace these services (i.e. the natural asset is characterised by limited substitution possibilities), then the marginal shadow value of the service is likely to rise all the more rapidly as the asset is increasingly degraded or converted (see, for example, Gerlagh and van der Zwaan 2002).[67] Hoel and Sterner (2007) and Sterner and Persson (2008) show how the marginal shadow value of a scarce environmental service might increase over time.[68]

[67] Gerlagh and van der Zwaan (2002) look at the case where individuals have a very strong preference for natural assets rather than non-substitutability per se (i.e. quasi-lexicographic preferences; see Rosenberger et al. 2003). The implications of this assumption, however, are that liquidating a natural asset beyond some threshold plausibly lowers the maximum level that future wellbeing can take.

[68] Hoel and Sterner (2007) use a constant elasticity of substitution (CES) system to describe how wellbeing derived from both 'environmental quality' and 'all other consumption goods' are substitutable for one another. As they show, this has implications for the way in which changes in the marginal shadow value of an ecosystem service are calculated. This change depends on the interaction of three elements. First, faster growth in other consumption will increase marginal shadow values (because people are generally wealthier and so will be willing to pay more for environmental quality). Secondly, if environmental quality is decreasing then this the increase in shadow values will be greater (reflecting the fact that this good is becoming scarcer). Lastly, a low elasticity of substitution implies a larger increase in the relative price of environmental quality because the wellbeing provided is not easily compensated for by increases in consumption of other goods. Sterner and Persson (2008) use this framework to show how estimates of future climate change damages are sensitive to changes in assumptions about substitutability.

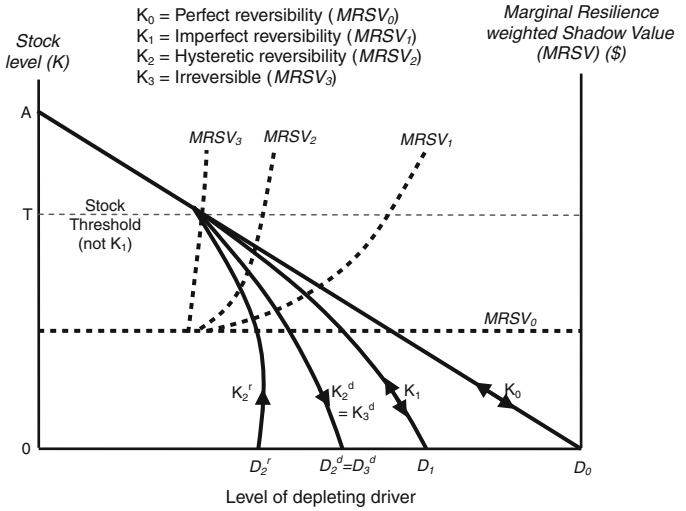


Fig. 6 Trade-off relationships between ecosystem stocks and depletion drivers with corresponding marginal resilience *weighted shadow values* (MRSV). (For simplicity we assume that these stocks are not substitutes for each other and stock level in all other respects these stocks are identical, yielding equal utility and shadow value where they are not depleted)

To calculate this, a number of assumptions must be made. Most notably, a judgement needs to be arrived at about the ease (or difficulty) with which particular natural assets can be replaced (the 'elasticity of substitution' between this and other assets). The less the potential for such replacement (i.e. lower elasticity of substitution), then the faster the increase in the marginal shadow value of a service as the underlying natural asset becomes scarcer. At the extreme certain natural assets may be 'critical' (Turner and Pearce 1993; Turner et al. 1994; Ekins et al. 2003) in that there are no substitution possibilities available. Such 'critical natural capital' requires particular care as discussed subsequently with respect to safe minimum standards.

8.2 Resilience Value

An innovative approach to the problem of assessing sustainability is proposed by Mäler et al. (2009) and Mäler (2008) who considers the ability of an ecosystem to withstand stresses and shocks and so continue to provide services. Mäler et al. propose treating this ecological 'resilience' as a stock with a distinct asset value which can be degraded or enhanced over time.

Walker et al. (2010) provides an empirical application of Mäler's resilience approach examining the value to agriculture in South-East Australia of maintaining a saline free water table (mainly through farmers cutting down trees to expand agriculture). Here agricultural expansion represents a driver depleting the stock of non-salinated soils (measured as the depth of soils for which saline intrusion is not a problem). As this depletion driver is increased so the stock of ecological resilience falls. As the depleting process itself may generate benefits (here agricultural produce) there is a trade-off to be assessed between the benefits of depletion and the fact that losses of resilience may need to be reversed if stocks fall below some threshold level.

The relationship between resilience stocks and depletion drivers can take many forms, the most simple of which is illustrated by function K_0 in Fig. 6. Here the trade off between

resilience stock and depletion driver is linear and perfectly reversible (i.e. there is no threshold). This implies that a given unit change in the depletion drive (either increase or decrease) results in a constant response (negative or positive) in resilience stocks. The stock of resilience available at any given time will generate a marginal shadow value. We can weight this value by the difficulty of reversing the level of prior depletion and can refer to this as the marginal resilience weighted shadow value (MRSV).[69] In the case illustrated by function K_0, as the relationship between depletion and stocks is perfectly reversible so the MRSV is a constant irrespective of the level of depletion (shown as $MRSV_0$).

While function K_0 exhibits a constant trade off between depletion drivers and resilience stocks irrespective of the level of the latter, this need not be the case. Other resources may exhibit threshold effects in that once stocks are depleted to some ecologically relevant level then further increases in the depletion driver result in an accelerating rate of stock reductions. This is reflected in the concavity of the stock depletion curve for resource K_1. This elevation in the rate of depletion once the threshold is breeched is reflected in the resilience weighted shadow value of K_1 shown by line $MRSV_1$ which is common with $MRSV_0$ when stocks are above the threshold but then rises above the latter reflecting the elevated rate of depletion.

Despite its threshold effect, stock K_1 still exhibits reversibility along the depletion path. However, this need not be the case. Stock K_2 also exhibits a threshold effect with depletion path K_2^d. However, once complete depletion is arrived at, the reversibility path (K_2^r) departs from the depletion path. Specifically, the level of the depletion driver has to be significantly reduced before the stock of resilience can begin to recover;[70] a phenomena known as hysteresis. Such a case is discussed by Walker et al. (2010) who notes that as agriculture in south-east Australia expands it cuts down trees which, due to the hydrological condition of the area, results in an increasing level of ground water salinisation. However, in order to reverse this situation farmers actually need to plant more trees than they originally felled because the elevated salinisation stunts the growth of newly planted trees. The additional costs implicit in having to further reduce depletion drivers (such as agriculture) in order to replenish ecological stocks (here the levels of non-salinised water) implies an elevated post-threshold shadow value of resilience, as illustrated in $MRSV_2$.

Both the imperfect and hysteretic reversibility cases imply that depletion of resources which have ecological thresholds imposes additional welfare losses above those associated with the perfect reversibility case. These elevated shadow values reach a maximum when we consider cases such as resource K_3. This follows depletion path K_3^d but once severe depletion is arrived at, reductions in the depletion driver fail to induce recovery in the stock (e.g. extending the Walker et al. case, imagine a situation where induced salinisation meant that all newly planted trees died and restoration was effectively impossible). This results in the extreme shadow value function $MRSV_3$. Note that this and all of the shadow value curves will become even steeper with increasing levels of risk aversion.

The resilience approach to sustainability provides a useful tool for economic analysis of ecosystem assessments. However, as Walker et al., acknowledge, it requires high degrees of knowledge and data availability, particularly when stocks with thresholds are influenced by multiple interacting variables. Therefore, we conclude our review of approaches to sustainability by considering a cruder but simpler method; safe minimum standards.

[69] Walker et al. 2010 define the MRSV as the probability of a change into an imperfectly reversible state (i.e. the chance of breeching a threshold) multiplied by the damage that such a change will cause.

[70] Such resource extraction effects are mirrored by the phenomena of pollution discharges which exceed the ambient environment's capacity to safely assimilate such emissions leading to a long run reduction in that capacity (Pearce 1976; Turner et al. 1994).

8.3 Safe Minimum Standards

A safe minimum standard (SMS) can be thought of as a precautionary approach to the management of a natural asset (Ciriacy-Wantrup 1952; Bishop 1978).[71] Here conventional economic decision making prevails unless a threshold threat is identified at which point the onus of proof shifts away from assuming that development is justified unless the costs to the environment do not justify proceeding, to a presumption that conservation is the right option unless the sacrifice (i.e. the opportunity costs) that it entails is intolerably high. Farmer and Randall (1998) and Randall (2007) argue that the appeal of the SMS lies in it being an approach that may well command broad moral consensus for making decisions. Allied to this is the argument that, in a situation of considerable uncertainty regarding underlying ecological processes, thresholds and economic values (to which we return subsequently), the SMS approach provides a safety-first approach to ensuring the future sustainability of human society through maintenance of the ecosystem stocks and services upon which it is reliant.

This thinking about SMS appears to have influenced conservation policy internationally (Berrens 2001; Pearce 2004), including in the US (e.g. the Endangered Species Act, ESA), Europe (e.g. the EU Birds and Habitats Directives, the Marine Strategy Directive, the Water Framework Directive, the European Landscape Convention, etc.) and the UK (e.g. Public Service Agreement 28, the Environmental Protection Act, Integrated Pollution Prevention and Control, Water Resources and Water Acts, Natural Environment and Rural Communities Act, Wildlife and Countryside Act, Forestry Act and others). Taking a US example, Berrens et al. (1998) investigate the implications of adopting these standards in case studies of two river systems. Common to both cases was the designation of areas of critical habitat for fish species identified as being endangered under the ESA. The sacrifice that this entailed involved restricting human uses of the rivers (by agricultural, industrial and household sectors). One question is where exactly to draw the line as regards when such sacrifice entailed by defending the SMS is too high. This level of 'tolerance' might be determined by the political process, by reference to some notional benchmark (such as an income loss indicators; see Berrens et al. 1998) or if the forgone development causes severe hardship or poverty (Pearce et al. 2006; Randall 2007). However, such considerations need to be balanced against the long-term consequences of ignoring threshold effects. Rockström et al. (2009) discuss numerous assets which exhibit threshold effects[72] and consider the potential for multiple global triggers of threshold effects, or 'tipping points' (Rockström et al. 2009). They argue that piecemeal abatement strategies are unlikely to be sufficiently effective in the face of such stressors and propose the adoption of a precautionary SMS to maintain the planet within a "safe operating space". For example, in the case of biodiversity loss they advocate a boundary of ten times the background rates of extinction. Because of the many gaps in our knowledge this boundary should be considered as preliminary, but they suggest that the current rate of species loss (100–1,000 times more than what could be considered natural) will lead to significant reductions in ecosystem resilience.[73]

[71] See also Gren et al. (1994) and Turner et al. (2003) who also argue for a pre-existing 'primary', 'glue' or 'infrastructure' value arising from an intact state for nature.

[72] Rockström et al. (2009) identify nine threshold effects: climate change; rate of biodiversity loss; interference with the nitrogen and phosphorus cycles; stratospheric ozone depletion; ocean acidification; global freshwater use; change in land use; chemical pollution and atmospheric aerosol loading. They argue that the first three of these have already reached the threshold zone.

[73] In practice the relationship between biodiversity and ecosystem services is complicated and it is not clear that a single threshold has meaning nor that it is best measured in terms of species loss (Díaz et al. 2007). Certain ecosystem services, for example erosion control or carbon sequestration, may be relatively insensitive

The implementation of any global safe minimum standards strategy will be controversial and will require concerted and targeted natural and social science research efforts to underpin it. Globalisation has resulted in a rate and extent of economic activity sufficient to pressurise a range of earth processes simultaneously. Addressing this is likely to require a radical overhaul of the governance processes controlling international trade and finance and resource exploitation etc. (Norgaard 2009). In the interim, recent work by Lenton et al. (2008) has proposed the use of early warning systems using trend analysis and predictive modelling to identify systems that are likely to cross such tipping points.

9 Concluding Comments

There is an expanding literature and interest in the concept of ecosystem services as an underpinning guide for decision making. The paper has attempted to contribute to this literature by provide an overview of issues arising from the economic analysis of such ecosystem service assessments. The paper opens by considering a single assessment period during which ecosystem assets are used in a sustainable manner. With these assumptions in place we begin our discussions by developing a terminology for future applications and employ this to focus upon the flow of ecosystems services and their role in generating welfare bearing goods and resultant benefits. We consider the distinction between sustainability analyses of historic trends and assessments of alternative policy options through scenario analyses. Economic analysis of the role and value of ecosystem services begins through isolating their contribution to welfare bearing goods. This contribution is then valued through the application of a range of methods and techniques ranging from adjustments of market prices to the measurement and valuation of preferences for non-market goods.

We illustrate the preceding principles though a case study focusing upon the economic value of the ecosystem services and associated goods arising from land use change. As part of this illustration we consider the central concepts of marginal analysis and substitution effects within economic assessments and relax our single period assumption to discuss the incorporation of multiple period costs and benefits within the assessment of shadow values; a process which introduces the concept of discounting. Finally we relax our assumption that assets are necessarily being managed in a sustainable manner. We consider approaches to incorporating such problems within economic analysis focussing upon the increase in real shadow values which this may induce, the explicit incorporation of resilience values within such analyses and the potential for adopting safe minimum standards as a solution to the depletion of ecosystem assets and the avoidance of stock threshold and tipping point effects.

Footnote 73 continued
to biodiversity and rely more strongly on the extent of structural biomass. On the other hand, ecosystem processes such as productivity and decomposition rates always show a positive relationship with biodiversity measured as functional trait diversity, although the relationship generally saturates at some point. However, empirical and theoretical studies indicate that in the case of long term environmental change or a constantly fluctuating environment, and to realise the full insurance value of biodiversity, much less loss of biodiversity can be tolerated (Yachi and Loreau 1999; Loreau et al. 2001). Finally, in the case of ecosystem service such as the non-use values held for wild species diversity and for potential benefits from bio-prospecting, any loss of diversity in species or genetic terms represents a significant loss and most commentators would not regard rates of loss comparable to those currently witnessed as acceptable. Therefore the role of biodiversity, while needing to be better understood, is probably best evaluated in context as was done for a recent European assessment (EASAC 2009) and as is planned for the UK NEA.

Given the wealth of issues which have to be embraced within an economic analysis of ecosystem services, our discussion so far has been deliberately uncritical in the main. However, we conclude by briefly touching upon some of the major problems facing such undertakings.

A major issue concerns the substantial gaps in natural science knowledge regarding the processes and interactions through which ecosystem services and provided and maintained. As Farley (2008) notes, the most serious problem facing ecosystem service assessments "is the ignorance and uncertainty concerning ecosystem function" (p1). The past century of natural science research has seen the development of a disciplinary expertise which, although vital in facilitating a deep understanding of individual systems, has inadvertently discouraged the cross-disciplinary cooperation and holism necessary to comprehend the interplay of ecological systems. While recent years have witnessed a marked reversal of this trend with interdisciplinary research being brought to the fore, it is still the case that interactions between systems are relatively poorly understood (Klein 1990, 2004; Lattuca 2001; Rhoten 2004). Furthermore, even this knowledge is stretched to its limits when we consider system interactions within the context of a general and unprecedented stressor such as global climate change. Natural science is being asked to predict into areas beyond the range of prior data and not surprisingly this raises new uncertainties, yet such problems have to be tackled if we are to employ the ecosystem service approach as a useful aid to practical policy and decision making.

Similar concerns and gaps exist within the economic literature. Interdisciplinary research, combining economics and other social sciences with the natural sciences, has advanced significantly over recent decades yet remains a relatively immature area of study conducted by small groups. Not surprisingly then, most natural scientists have little or no experience of the informational requirements for economic analysis (e.g. provision change assessments for economic valuation purposes; damage and threshold analyses for resilience measures, etc.). Furthermore, even within the economics camp, the continuing emphasis upon conventional economic research conducted as a single discipline means that the efforts of environmental and resource economists remain a minority branch of research. For example, even after four decades of research into the valuation of non-market goods such as those provided by the natural environment, the available stock of knowledge remains inadequate given the complexity of the demands upon it. The development of value transfer methods enhances the usefulness of the available literature but cannot make up for an absence of high quality valuation studies regarding key questions such as the impact upon marginal shadow values of depleting ecological stocks and accompanying threshold effects. In the main this situation reflects the very small base from which such research started and its initially slow growth. However, it also indicates an inherent aversion to strategic investment in wider ecosystem service valuation with the focus instead being dictated by short term investment and policy priorities concentrating upon single resources. This results in a degree of uncertainty regarding the value of different provision options which is far from satisfactory and has led some to call for decisions to be made without resort to valuation (e.g. Vatn and Bromley 1994). However, an alternative strategy would be to infer value relationships based upon economic theory and related intuition. For example, while we may have sparse data on the value of a given good, we can nonetheless draw upon theory to infer likely shapes for marginal benefits curves, incorporating diminishing or increasing marginal values as provision levels rise or fall. Furthermore, this issue is overtly acknowledged within the remit of the UK NEA and indeed one of its aims is to identify knowledge gaps for the purposes of directing future research.

Other concerns are well documented yet equally important. An obvious issue is the topic of discounting. Within the conventional paradigm it is still true that, if the bulk of society is as

some would argue, self obsessed and relatively uninterested in others, including future generations, then a discounting regime which reflects that will lead to economic analyses which maximise present generation benefits even if outcomes are ecologically unsustainable. We have reviewed potential solutions such as precautionary safe minimum standards or ethics based discounting, yet to date these are not the generally accepted norm.

A further well established concern is that economic valuations generally reflect the current distribution of income with those with higher ability to pay being better able to reflect their preferences through higher willingness to pay. Arguably this is not a concern given that most decisions are predicated on existing income distributions and consequent demand. However, from an economic welfare perspective such situations may very well not optimise social wellbeing. In a manner similar to the discounting issue, solutions are available. For example, within the Treasury (2003) official UK decision making guidelines higher weighting is accorded to lower income groups because of the greater value of income gains to the latter.[74] It should be noted however that, despite the official acceptability of such approaches, they are rarely implemented within economic analyses of ecosystem services. The UK NEA seeks to remedy this situation by undertaking such adjustments where data permits.

In conclusion then, we have attempted throughout the main body of this paper to provide an overview of the variety of issues raised by the application of economic analysis techniques to ecosystem service assessments. As noted, these are complex analyses and as we acknowledge above, the issues raised are not settled and the available literature less than complete. Nevertheless the importance of incorporating ecosystems services within policy appraisal and decision making techniques cannot be understated; it is literally a life and death issue for the future of human society and the difficulties raised cannot be shirked or the task ignored. Many would argue that there is no more important issue facing the natural and social sciences today.

Open Access This article is distributed under the terms of the Creative Commons Attribution Noncommercial License which permits any noncommercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

References

Adamowicz W, Louviere J, Williams M (1994) Combining revealed and stated preference methods for valuing environmental amenities. J Environ Econ Manage 26(3):271–292

Adger WN, Brown K, Shiel RS, Whitby MC (1992) Carbon dynamics of land use in Great Britain. J Environ Manage 36:117–133

Arnade C, Kelch D (2007) Estimation of area elasticities from a standard profit function. Am J Agric Econ 89:727–737

Arrow KJ, Dasgupta P, Goulder LH, Mumford K, Oleson K (2007) China, the US, and sustainability: perspectives based on comprehensive wealth. Stanford working paper

Ayres RU (2007) On the practical limits to substitution. Ecol Econ 61(1):115–128. doi:10.1016/j.ecolecon.2006.02.011

Badola R, Hussain SA (2005) Valuing ecosystem functions: an empirical study on the storm protection function of Bhitarkanika mangrove ecosystem, India. Environ Conserv 32(1):85–92

Balmford A, Bruner A, Cooper P, Constanza R, Farber S, Green RE, Jenkins M, Jefferis P, Jessamy V, Madden J, Munro K, Myers N, Naeem S, Paavola J, Rayment M, Rosendo S, Roughgarden J, Trumper K, Turner RK (2002) Economic reasons for conserving wild nature. Science 297:950–953. doi:10.1126/science.1073947

[74] See also the income reweighting approach to implementing cost-benefit analysis discussed by Pearce (1983).

Banzhaf HS, Burtraw D, Evans D, Krupnick A (2006) Valuation of natural resource improvements in the adirondacks. Land Econ 82(3):445–464
Barbier EB (1994) Valuing environmental functions: tropical wetlands. Land Econ 70:155–173
Barbier EB (2000) Valuing the environment as input: applications to mangrove-fishery linkages. Ecol Econ 35:47–61
Barbier EB (2003) Habitat-fishery linkages and mangrove loss in Thailand. Contemp Econ Policy 21(1):59–77
Barbier EB (2007) Valuing ecosystem services as productive inputs. Econ Policy 22:177–229
Barbier EB (2009) Ecosystems as natural assets. Found Trends Microecon 4(8):611–681
Barbier EB, Koch EW, Silliman BR, Hacker SD, Wolanski E, Primavera J, Granek EF, Polasky S, Aswani S, Cramer LA, Stoms DM, Kennedy CJ, Bael D, Kappel CV, Perillo GM, Reed DJ (2008) Coastal ecosystem-based management with nonlinear ecological functions and values. Science 319:321–323
Barbier EB, Markandya A, Pearce DW (1990) Environmental sustainability and cost—benefit analysis. Environ Plann A 22(9):1259–1266
Bartelmus P (2001) Accounting for sustainability: greening the national accounts. In: Tolba MK (ed) Our fragile world, forerunner to the encyclopedia of life support systems II. Eolss Publishers, Oxford, pp 1721–1735
Bartelmus P (2008) Quantitative eco–nomics, how sustainable are our economies? Springer Science and Business Media, Secaucus NJ and Heidelberg
Barton DN (2002) The transferability of benefit transfer: contingent valuation of water quality improvements in Costa Rica. Ecol Econ 42:147–164
Bateman IJ (2007) Valuing preferences regarding environmental change. In: Pretty J, Ball A, Benton T, Guivant J, Lee D, Orr D, Pfeffer M, Ward H (eds) The SAGE handbook of environment and society. Sage, London, pp 155–171
Bateman IJ (2009a) Economic analysis for ecosystem assessments: application to the UK national ecosystem assessment (NEA). CSERGE working paper EDM 09–12, Centre for Social and Economic Research on the Global Environment, University of East Anglia
Bateman IJ (2009b) Bringing the real world into economic analyses of land use value: incorporating spatial complexity. Land Use Policy 26:S30–S42. doi:10.1016/j.landusepol.2009.09.010
Bateman IJ, Jones AP (2003) Contrasting conventional with multi-level modelling approaches to meta-analysis: an illustration using UK woodland recreation values. Land Econ 79(2):235–258
Bateman IJ, Carson RT, Day B, Hanemann WM, Hanley N, Hett T, Jones-Lee M, Loomes G, Mourato S, Özdemiroğlu E, Pearce DW, Sugden R, Swanson J (2002a) Economic valuation with stated preference techniques: a manual. Edward Elgar Publishing, Cheltenham
Bateman IJ, Jones AP, Lovett AA, Lake I, Day BH (2002b) Applying geographical information systems (GIS) to environmental and resource economics. Environ Resour Econ 22(1–2):219–269. doi:10.1023/A:1015575214292
Bateman IJ, Lovett AA, Brainard JS (2003) Applied environmental economics: a GIS approach to cost-benefit analysis. Cambridge University Press, Cambridge
Bateman IJ, Day BH, Georgiou S, Lake I (2006) The aggregation of environmental benefit values: welfare measures, distance decay and total WTP. Ecol Econ 60(2):450–460. doi:10.1016/j.ecolecon.2006.04.003
Bateman IJ, Burgess D, Hutchinson WG, Matthews DI (2008a) Contrasting NOAA guidelines with learning design contingent valuation (LDCV): preference learning versus coherent arbitrariness. J Environ Econ Manage 55:127–141. http://dx.doi.org/10.1016/j.jeem.2007.08.003
Bateman IJ, Munro A, Poe GL (2008b) Asymmetric dominance effects in choice experiments and contingent valuation. Land Econ 84:115–127. http://le.uwpress.org/cgi/reprint/84/1/115
Bateman IJ, Day BH, Dupont D, Georgiou S (2009a) Procedural invariance testing of the one-and-one-half-bound dichotomous choice elicitation method. Rev Econ Stat 91(4):806–820. doi:10.1162/rest.91.4.806
Bateman IJ, Day BH, Jones AP, Jude S (2009b) Reducing gains/loss asymmetry: a virtual reality choice experiment (VRCE) valuing land use change. J Environ Econ Manag 58:106–118. doi:10.1016/j.jeem.2008.05.003
Bateman IJ, Coombes E, Fisher B, Fitzherbert E, Glew DW, Naidoo R (2009c) Saving sumatra's species: combining economics and ecology to define an efficient and self-sustaining program for inducing conservation within oil palm plantations. In: CSERGE working paper EDM-2009-03, centre for social and economic research on the global environment. University of East Anglia
Bateman IJ, Brouwer R, Ferrini S, Schaafsma M, Barton DN, Dubgaard A, Hasler B, Hime S, Liekens I, Navrud S, De Nocker L, Ščeponavičiūtė R, Semėnienė D (2009d) Making benefit transfers work: deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the

non-market benefits of water quality improvements across Europe. In: CSERGE working paper EDM 09–10, centre for social and economic research on the global environment. University of East Anglia
Bateman IJ, Fisher B, Fitzherbert E, Glew D, Naidoo R (2010a) Tigers, markets and palm oil: market potential for conservation. Oryx 44(2):230–234. doi:10.1017/S0030605309990901
Bateman IJ, Brouwer R, Cranford M, Hime S, Ozdemiroglu E, Phang Z, Provins A (2010b) Valuing environmental impacts: practical guidelines for the use of value transfer in policy and project appraisal, main report to the department for environment, food and rural Affairs. Economics for the Environment Consultancy (eftec), London. Published at http://www.defra.gov.uk/environment/policy/natural-environ/documents/guidelines.pdf
Bateman IJ, Binner A, Coombes E, Day BH, Ferrini S, Fezzi C, Hutchins M, Posen P (2010c) Integrated and spatially explicit modelling of the economic value of complex environmental change and its knock-on effects. In: Presented at the 4th world congress of environmental and resource economists (WCERE2010), Montreal, Canada, 28 June–2 July 2010
Beckerman W, Hepburn C (2007) Ethics of the discount rate in the stern review on the economics of climate change. World Econ 8:187–210
Bergland O, Magnussen K, Navrud S (1995) Benefit transfer: testing for accuracy and reliability. In: Florax RJGM, Nijkamp P, Willis KG (eds) (2002) comparative environmental economic assessment. Edward Elgar, UK
Berrens RP (2001) The safe minimum standard of conservation and endangered species: a review. Environ Conserv 28:104–116
Berrens RP, Brookshire DS, McKee M, Schmidt C (1998) Implementing the safe minimum standard approach: two case studies from the US endangered species act. Land Econ 74(2):147–161
Bishop R (1978) Endangered species and uncertainty: the economics of a safe minimum standard. Am J Agric Econ 60:10–18
Bockstael NE, McConnell KE (2006) Environmental and resource valuation with revealed preferences: a theoretical guide to empirical models, the economics of non-market goods and services, vol 7. Springer, Dordrecht
Bockstael NE, Freeman AM III, Kopp RJ, Portney PR, Smith VK (2000) On measuring economic values for nature. Environ Sci Technol 34:1384–1389
Boscolo M, Vincent JR (2003) Nonconvexities in the production of timber, biodiversity, and carbon sequestration. J Envir Econ Manage 46:251–268
Boyd J, Banzhaf S (2007) What are ecosystem services? The need for standardized environmental accounting units. Ecol Econ 63:616–626
Boyd JW, Krupnick A (2009) The definition and choice of environmental commodities for nonmarket valuation. RFF Discussion Paper 09–35. Resources for the Future, Washington
Brander LM, Florax RJGM, Vermaat JE (2006) The empirics of wetland valuation: a comprehensive summary and a meta-analysis of the literature. Envir Resour Econ 33:223–250. doi:10.1007/s10640-005-3104-4
Brouwer R (2000) Environmental value transfer: state of the art and future prospects. Ecol Econ 32:137–152
Brouwer R, Bateman IJ (2005) Benefits transfer of willingness to pay estimates and functions for health-risk reductions: a cross-country study. J Health Econ 24: 591–611. doi:10.1016/j.jealeco.2004.09.004
Brouwer R, Langford IH, Bateman IJ, Turner RK (1999) A meta-analysis of wetland contingent valuation studies. Reg Envir Chang 1(1):47–57
Brouwer R, Spanninks FA (1999) The validity of environmental benefits transfer: further empirical testing. Environ Resour Econ 14(1):95–117
Cairns RD (2002) Green accounting using imperfect, current prices. Environ Dev Econ 7(2):207–214
Carlsson F, Frykblom P, Lijenstolpe C (2003) Valuing wetland attributes: an application of choice experiments. Ecol Econ 47:95–103
Carpenter SR, Mooney HA, Agard J, Capistrano D, DeFries RS, Diaz S, Dietz T, Duraiappah AK, Oteng-Yeboah A, Pereira HM, Perrings C, Reid WV, Sarukhan J, Scholes RJ, Whyte A (2009) Science for managing ecosystem services: beyond the millennium ecosystem assessment. Proc Natl Acad Sci USA 106:1305–1312
Carson RT, Flores NE, Martin KM, Wright JL (1996) Contingent valuation and revealed preference methodologies: comparing the estimates for quasi-public goods. Land Econ 72:80–99
Carson RC, Mitchell R, Hanemann WM, Kopp R, Presser S, Ruud P (2003) Contingent valuation and lost passive use: damages from the exxon valdez oil spill. Environ Resour Econ 25(3):257–286
Champ PA, Boyle K, Brown TC (eds) (2003) A primer on non-market valuation, the economics of non-market goods and services, vol 3. Kluwer Academic Press, Dordrecht
Chapin FS III, Zavaleta ES, Valerie TE, Naylor RL, Vitousek PM, Reynolds HL, Hooper DU, Lavorel S, Sala OE, Hobbie SE, Mack MC, Díaz S (2000) Consequences of changing biodiversity. Nature 405:234–242. doi:10.1038/35012241

Chiang AC (1984) Fundamental methods of mathematical economics. McGraw Hill, New York
Chichilnisky G, Heal GM (1998) Economic returns from the biosphere. Nature 391:629–630
Christie M, Hanley N, Warren J, Murphy K, Wright R, Hyde T (2006) Valuing the diversity of biodiversity. Ecol Econ 58:304–317
Ciriacy-Wantrup Sv (1952) Resource conservation: economics and policies. University of California Press, Berkeley
Considine TJ, Larson DF (2006) The environment as a factor of production. J Environ Econ Manage 52: 645–662
Conventionon on Biological Diversity (CBD) (2006) Global biodiversity outlook 2. Secretariat of the Convention on Biological Diversity, Montreal
Costanza R, d'Arge R, Groot Rd, Farber S, Grasso M, Hannon B, Limburg K, Naeem S, O'Neill RV, Paruelo J, Raskin RG, Sutton P, Belt Mvd (1997) The value of the world's ecosystem services and natural capital. Nature 387:253–260
Costanza R, Pérez-Maqueo O, Martinez ML, Sutton P, Anderson SJ, Mulder K (2008) The value of coastal wetlands for hurricane protection. Ambio 37:241–248
Daily G (1997) Nature's services: societal dependence on natural ecosystems. Island Press, Washington
Dasgupta P (2001) Human wellbeing and the natural environment. Oxford University Press, Oxford
Dasgupta P (2007) Comments on the stern review's economics of climate change. Natl Inst Econ Rev 199:4–7
Dasgupta P (2009) The welfare economic theory of green national accounts. Environ Resour Econ 42(1):3–38. doi:10.1007/s10640-008-9223-y
Dasgupta P, Mäler K-G (2000) Net national product, wealth, and social well-being. Environ Dev Econ 5(1):69–93
Day BH, Bateman IJ, Lake I (2007) Beyond implicit prices: recovering theoretically consistent and transferable values for noise avoidance from a hedonic property price model. Environ Resour Econ 37(1):211–232. doi:10.1007/s10640-007-9121-8
De Groot RS, Wilson MA, Boumans RMJ (2002) A typology for the classification, description and valuation of ecosystem functions, goods and services, special issue on the dynamics and value of ecosystem services: integrating economic and ecological perspectives. Ecol Econ 41:393–408
Defra (2009) Farm accounts in England: results of the farm business survey. Department for Environment Food & Rural Affairs, London
Desvousges WH, Naughton MC, Parsons GR (1992) Benefit transfer: conceptual problems in estimating water quality benefits using existing studies. Water Resour Res 28(3):675–683
Díaz S, Lavorel S, de Bello F, Quétier F, Grigulis K, Robson M (2007) Incorporating plant functional diversity effects in ecosystem service assessments. Proc Natl Acad Sci USA 104:20684–20689
Dietz S, Hepburn C (2010) On non-marginal cost-benefit analysis. Grantham Research Institute on Climate Change and the Environment, London School of Economics and Political Science, London
Dietz S, Anderson D, Stern N, Taylor C, Zenghelis D (2007) Right for the right reasons: a final rejoinder on the stern review. World Econ 8:229–258
Downing M, Ozuna T (1996) Testing the reliability of the benefit function transfer approach. J Environ Econ Manage 30:316–322
EASAC (2009) Ecosystem services and biodiversity in Europe. EASAC policy report 09 European Academies Science Advisory Council, London
Eaton MA, Brown AF, Noble DG, Musgrove AJ, Hearn R, Aebischer NJ, Gibbons DW, Evans A, Gregory RD (2009) Birds of conservation concern 3: the population status of birds in the United Kingdom, Channel Islands and the Isle of man. Br Birds 102:296–341
Egan K, Herriges JA (2006) Multivariate count data regression models with individual panel data from an on-site sample. J Environ Econ Manage 52(2):567–581
Ekins P, Folke C, De Groot R (2003) Identifying critical natural capital. Ecol Econ 44(2–3):159–163
Ellis GM, Fisher AC (1987) Valuing the environment as input. J Environ Manage 25:149–156
Environment Agency (2001) Water resources for the future: a strategy for England and Wales. UK Environment Agency, London
Farley J (2008) Valuing natural capital: the limits of marginal valuation in complex systems. In: Presented at economics and conservation in the tropics: a strategic dialogue, the conservation strategy fund, resources for the future, and the gordon and betty moore foundation, San Francisco, January 31–February 1 2008
Farmer MC, Randall A (1998) The rationality of a safe minimum standard. Land Econ 74:287–302
FBU (Farm Business Unit) (2009) Farm business survey in wales. University of Wales, Aberystwyth
Federal Register (1995) Federal guidance for the establishment, use and operation of mitigation banks. Fed Regist 60(228):58605–58614
Fezzi C, Bateman IJ (2009) Structural agricultural land use modelling. In: Presented at The XXVII international conference of agricultural economists, Beijing, China, 16th–22nd August 2009. Available at http://ageconsearch.umn.edu/handle/51423

Fezzi C, Bateman IJ (2010) The impact of climate change on agriculture: a panel ricardian analysis using GAMMs. In: Presented at envecon 2010: applied environmental economics conference, 12th March 2010. The Royal Society, London
Finnoff D, Tschirhart J (2008) Linking dynamic economic and ecological general equilibrium models. Resour Energy Econ 30:91–114
Fisher B, Turner RK (2008) Ecosystem services: classification for valuation. Biol Conserv 141(5):1167–1169
Fisher B, Turner RK, Zylstra M, Brouwer R, De Groot R, Farber S, Ferraro P, Green R, Hadley D, Harlow J, Jefferiss P, Kirkby C, Morling P, Mowatt S, Naidoo R, Paavola J, Strassburg B, Yu D, Balmford A (2008) Ecosystem services and economic theory: integration for policy-relevant research. Ecol Appl 18(8):2050–2067
Foresight Land Use Futures Project (2010) Final project report. The Government Office for Science, London
Freeman AM III (1991) Valuing environmental resources under alternative management regimes. Ecol Econ 3:247–256
Freeman AMIII (2003) The measurement of environmental and resource values: theory and methods, 2nd edn. Resources for the Future, Washington
Gerlagh R, van der Zwaan BCC (2002) Long-term substitutability between environmental and man-made goods. J Environ Econ Manage 44:329–345
Global Environment Facility (GEF) (1998) Valuing the global environment: actions and investments for a 21st century. GEF, Washington
Godoy R, Lubowski R, Markandya A (1993) A method for the economic valuation of non-timber forest products. Econ Bot 47:220–233
Gren I-M, Elofsson K, Jannke P (1997) Cost-effective nutrient reductions to the Baltic Sea. Environ Resour Econ 10:341–362
Gren I-M, Folke C, Turner RK, Bateman IJ (1994) Primary and secondary values of wetland ecosystems. Environ Resour Econ 4:55–74
Griffin RC, Montgomery JM, Rister ME (1987) Selecting functional form in production function analysis. West J Agric Econ 12(02):216–227
Groom B, Hepburn C, Koundouri P, Pearce DW (2005) Declining discount rates: the long and the short of it. Environ Resour Econ 33:445–493
Guyomard H, Baudry M, Carpentier A (1996) Estimating crop supply response in the presence of farm programmes: application to the CAP. Eur Rev Agric Econ 23:401–420
Haab TC, McConnell KE (2002) Valuing environmental and natural resources. Edward Elgar Publishing, Cheltenham
Haines-Young R, Potschin M, de Groot D, Kienast F, Bolliger J (2009) Towards a common international classification of ecosystem services (CICES) for integrated environmental and economic accounting (Draft V1). In: Report to the European environment agency for contract No: No. EEA/BSS/07/007, centre for environmental management. School of Geography, University of Nottingham
Hamilton K, Clemens M (1999) Genuine Saving in developing countries. World Bank Econ Rev 13(2):333–356
Hamilton K, Withagen C (2007) Savings growth and the path of utility. Can J Econ 40(2):703–713
Hamilton K, Ruta G (2009) Wealth accounting, exhaustible resources and social welfare, themed issue: advances in the theory and practice of environmental accounting. Environ Resour Econ 42(1):53–64
Hanley N, Barbier EB (2009) Pricing nature: cost-benefit analysis and environmental policy-making. Edward Elgar, London
Hanley N, MacMillan D, Patterson I, Wright R (2003) Economics and the design of nature conservation policy: a case study of wild goose conservation in scotland using choice experiments. Anim Conser 6:123–129
Hansen L, Hellerstein D (2007) The value of the reservoir services gained with soil conservation. Land Econ 83(3):285–301
Hargrove C (1992) Weak anthropocentric intrinsic value. Monist 75:183–207
Heal G (2000) Valuing ecosystem services. Ecosystems 3(1):24–30
Heal GM, Barbier EB, Boyle KJ, Covich AP, Gloss SP, Hershner CH, Hoehn JP, Pringle CM, Polasky S, Segerson K, Shrader-Frechette K (2005) Valuing ecosystem services: toward better environmental decision making. The National Academies Press, Washington
Hearne RR, Salinas J (2002) The use of choice experiments in the analysis of tourist preferences for ecotourism development in costa rica. J Environ Manage 65:153–163
Herriges JA, Kling CL, Phaneuf DJ (2004) What's the use? Welfare estimates from revealed preference models when weak complementarity does not hold. J Environ Econ Manage 47(1):53–68
Hoel M, Sterner T (2007) Discounting and relative prices. Clim Change 84:265–280
House of Commons Environmental Audit Committee (2007) Government response to the committee's first report of session 2006–07: the UN millennium ecosystem assessment. HC 848, The Stationery Office Limited, London

Howarth RB, Farber S (2002) Accounting for the value of ecosystem services. Ecol Econ 41:421–429
Hulme M, Jenkins GL, Lu X, Turnpenny JR, Mitchell TD, Jones RG, Lowe J, Murphy JM, Hassell D, Boorman P, McDonald R, Hill S (2002) Climate change scenarios for the United Kingdom: the UKCIP02 scientific report. Tyndall Centre for Climate Change Research, School of Environmental Sciences, University of East Anglia, Norwich, UK, p 120
Huybers T, Bennett J (2003) Environmental management and the competitivenes of nature-based tourism. Environ Resour Econ 24:213–233
IPCC (1996) Climate change 1995: economic and social dimensions of climate change. Cambridge University Press, Cambridge
IPCC (2000) Special report on emissions scenarios (SRES). In: Nakicenovic N, Swart R (eds) International panel on climate change. Cambridge University Press, Cambridge, p 612
Johnstone RJ, Duke JM (2009) Willingness to pay for land preservation across states and jurisdictional scale: implications for benefit transfer. Land Econ 85(2):217–237
Johnston RJ, Besedin EY, Iovanna R, Miller CJ, Wardwell RF, Ranson MH (2005) Systematic variation in willingness to pay for aquatic resource improvements and implications for benefit transfer: a meta-analysis. Can J Agric Econ 53:221–248
Johnston RJ, Ranson MH, Besedin EY, Helm EC (2006) What determines willingness to pay per fish? A meta-analysis of recreational fishing values. Mar Resour Econ 21:1–32
Kahneman D, Wakker P, Sarin R (1997) Back to Bentham? Exploration of experienced utility. Q J Econ 112:373–405
Kaiser B, Roumasset J (2002) Valuing indirect ecosystem services: the case of tropical watersheds. Environ Dev Econ 7:701–714
Kanninen B (ed) (2006) Valuing environmental amenities using stated choice studies: a common sense approach to theory and practice, the economics of non-market goods and services, vol 8. Springer, Dordrecht
Kaval P (2010) A summary of ecosystem service economic valuation methods and recommendations for future studies. Working paper series #10Department of Economics, University of Waikato
Kim S-H, Dixon JA (1986) Economic valuation of environmental quality aspects of upland agricultural projects in Korea. In: Dixon JA, Hufschmidt MM (eds) Economic valuation techniques for the environment: a case study workbook. Johns Hopkins Press, Baltiimore
Klein JT (1990) Interdisciplinarity: history, theory, and practice. Wayne State University Press, Detroit
Klein JT (2004) Prospectus for transdisciplinarity. Futures 36:515–526
Knowler D, Barbier EB (2005) Managing the Black Sea anchovy fishery with nutrient enrichment and a biological invader. Mar Resour Econ 20:263–285
Knowler D, Barbier EB, Strand I (2001) An open-access model of fisheries and nutrient enrichment in the Black Sea. Mar Resour Econ 16:195–217
Koziell I (2001) Diversity not adversity: sustaining livelihoods with biodiversity. International Institute for Environment and Development and Department for International Development (DFID). England
Krutilla J, Fisher AC (1974) The economics of natural environments. Johns Hopkins University Press, Baltimore
Lattuca L (2001) Creativity interdisciplinarity: interdisciplinary research and teaching among college and university faculty. Vanderbilt University Press, Nashville
Layard R (2005) Happiness: lessons from a new science. Penguin Books, London
Lenton TM, Held H, Kriegler E, Hall JW, Lucht W, Rahmstorf S, Schellnhuber HJ (2008) Tipping elements in the Earth's climate system. PNAS 105:1786–1793
Leon-Gonzalez R, Scarpa R (2008) Improving multi-site benefit functions via Bayesian model averaging: a new approach to benefit transfer. J Environ Econ Manage 56(1):50–68
Lindhjem HL, Navrud HL (2008) How reliable are meta-analyses for international benefit transfers? Ecol Econ 66:425–435
Liu S, Costanza R, Farber S, Troy A (2010) Valuing ecosystem services. Ann N Y Acad Sci 1185:54–78. doi:10.1111/j.1749-6632.2009.05167.x
Loreau M, Naeem S, Inchausti P, Bengtsson J, Grime JP, Hector A, Hooper DU, Huston MA, Raffaelli D, Schmid B, Tilman D, Wardle DA (2001) Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294:804–808
Loreau M, Oteng-Yeboah A, Arroyo MTK, Babin D, Barbault R, Donoghue M, Gadgil M, Häuser C, Heip C, Larigauderie A, Ma K, Mace G, Mooney HA, Perrings C, Raven P, Sarukhan J, Schei P, Scholes RJ, Watson RT (2006) Diversity without representation. Nature 442:245–246. doi:10.1038/442245a
Mäler K-G (1974) Environmental economics: a theoretical inquiry. Resources for the Future, Baltimore
Mäler K-G (2008) Sustainable development and resilience in ecosystems. Environ Resour Econ 39(1):17–24

Mäler K-G, Aniyar S, Jansson Å (2008) Accounting for ecosystem services as a way to understand the requirements for sustainable development. PNAS 105(28):9501–9506. www.pnas.org/cgi/doi/10.1073/pnas.0708856105
Mäler K-G, Aniyar S, Jansson Å (2009) Accounting for ecosystems. Environ Resour Econ 42:39–51
Mansfield C, Phaneuf DJ, Johnson FR, Yang J-C, Beach R (2008) Preferences for public lands management under competing uses: the case of yellowstone national park. Land Econ 84(2):282–305
Mardle S, James C, Pipitone C, Kienzle M (2004) Bioeconomic interactions in an established fishing exclusion zone: the gulf of castellammare, NW Sicily. Nat Resour Model 17(4):393–447
Mendelsohn R, Nordhaus W, Shaw D (1994) The impact of global warming on agriculture: a ricardian analysis. Am Econ Rev 84:753–771
Millennium Ecosystem Assessment (MA) (2003) Ecosystems and human well-being: a framework for assessment. Island Press, Washington
Millennium Ecosystem Assessment (MA) (2004) Scenarios working group (unpublished report). www.millenniumassessment.org
Millennium Ecosystem Assessment (MA) (2005) Ecosystems and human well-being: a framework for assessment. Island Press, Washington DC
Moeltner K, Boyle KJ, Paterson RW (2007) Meta-analysis and benefit transfer for resource valuation-addressing classical challenges with Bayesian modeling. J Environ Econ Manage 53(2):250–269
Moore MR, Gollehon NR, Carey MB (1994) Multicrop production decisions in western irrigated agriculture: the role of water price. Am J Agric Econ 76:859–874
Murphy J, Sexton D, Jenkins G, Boorman P, Booth B, Brown K, Clark R, Collins M, Harris G, Kendon L (2009) Climate change projections. UK Climate Impacts Programme, Met Office Hadley Centre
Muthke T, Holm-Mueller K (2004) National and international benefit transfer testing with a rigorous test procedure. Environ Resour Econ 29:323–336
Naidoo R, Adamowicz WL (2005) Biodiversity and nature-based tourism at forest reserves in Uganda. Environ Dev Econ 10:159–178
Nalle DJ, Montgomery CA, Arthur JL, Polasky S, Schumaker NH (2004) Modeling joint production of wildlife and timber. J Environ Econ Manage 48(3):997–1017
Navrud S, Ready R (eds) (2007) Environmental value transfer: issues and methods. Springer, Dordrecht
Naylor R, Drew M (1998) Valuing mangrove resources in Kosrae, Micronesia. Environ Dev Econ 3:471–490
Nicholson E, Mace GM, Armsworth PR, Atkinson G, Buckle S, Clements T, Ewers RM, Fa JE, Gardner TA, Gibbons J, Grenyer R, Metcalfe R, Mourato S, Muûls M, Osborn D, Reuman DC, Watson C, Milner-Gulland EJ (2009) Priority research areas for ecosystem services in a changing world. J Appl Ecol 46:1139–1144
Norgaard R (2009) Ecosystem services: from eye-opening metaphor to complexity blinder. Ecol Econ 69(6):1219–1227. doi:10.1016/j.ecolecon.2009.11.009
Nordhaus WD (2007) A review of the stern review on the economics of climate change. J Econ Lit 45:686–702
Office of Science & Technology (OST) (1999) The use of scenarios in Foresight. An information document prepared for the OST by Professor David Stout. www.foresight.gov.uk
Office of Science & Technology (OST) (2003) Foresight futures 2020: revised scenarios and guidance. Department for Trade & Industry, UK, p 38. www.foresight.gov.uk
Oswald AJ, Wu S (2010) Objective confirmation of subjective measures of human well-Being: evidence from the USA. IZA Discussion paper no. 4695, Institute for the Study of Labor (IZA), Bonn
Othman J, Bennett J, Blamey R (2004) Environmental management and resource management options: a choice modelling experience in malaysia. Environ Dev Econ 9:803–824
Oude LA, Peerlings J (1996) Modelling the new EU cereals and oilseeds regime in the Netherlands. Eur Rev Agric Econ 23:161–178
Pachauri RK, Reisinger A (eds) (2007) Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change. IPCC, Geneva
Pagiola S, Ritter Kv, Bishop JT (2004) How much is an ecosystem worth? Assessing the economic value of conservation. The World Bank, Washington
Pascal U, Muradian R, Gómez-Baggethun E, Armsworth P, Brander L, Cornelissen H, Farley J, Loomes J, Martinez-López B, Pearson L, Perrings C, Polasky S, Verma M (2009) Valuation of ecosystems services: methodology and challenges. In: Report to review of the economics of ecosystems and biodiversity. European Commission/ UNEP/ BMU-Germany
Pearce DW (1976) The limits of CBA as a guide to environmental policy. Kyklos 29(1):97–112. doi:10.1111/j.1467-6435.1976.tb01962.x
Pearce DW (1983) Cost benefit analysis, 2nd edn. Macmillan, NY
Pearce DW (2004) Does european union Environmental policy pass a cost-benefit test? World Econ 5:115–138
Pearce DW, Atkinson G, Mourato S (2006) Cost-benefit analysis and the environment: recent developments. Organisation for Co-operation and Development (OECD), Paris

Pearce DW, Markandya A, Barbier EB (1989) Blueprint for a green economy. Earthcan, London
Perman R, Ma Y, McGilvray J (1996) Natural resource and environmental economics. Addison Wesley, New York
Pinnegar JK, Viner D, Hadley D, Dye S, Harris M, Berkout F, Simpson M (2006) Alternative future scenarios for marine ecosystems: technical report. Cefas Lowestoft, p 109
Prime Minister's Strategy Unit (PMSU) (2004) Net benefits: a sustainable and profitable future for UK fishing. Prime Minister's Strategy Unit, Cabinet Office, UK. p 168. http://www.number-10.gov.uk/su/fish/index.htm
Randall A (1994) A difficulty with the travel cost method. Land Econ 70:88–96
Randall A (2007) Benefit cost analysis and a safe minimum standard. In: Atkinson G, Dietz S, Neumayer E (eds) Handbook of sustainable development. Edward Elgar, Cheltenham
Ready R, Navrud R, Day B, Dubourg R, Machado F, Mourato S, Spaninks F, Vázquez Rodriquez MX (2004) Benefit transfer in Europe: how reliable are transfers between countries? Environ Resour Econ 29:67–82
Rhoten D (2004) Interdisciplinary research: trend or transition. Items Issues 5(1–2):6–11
Ricketts TH, Gretchen GC, Ehrlich PR, Michener CD (2004) Economic value of tropical forests to coffee production. Proc Nat Acad Sci 101(34):12579–12582
Rockström J, Steffen W, Noone K, Persson A, Chapin FSIII, Lambin EF, Lenton TM, Scheffer M, Folke C, Schellnhuber HJ, Nykvist B, de Wit CA, Hughes T, van der Leeuw S, Rodhe H, Sorlin S, Snyder PK, Costanza R, Svedin U, Falkenmark M, Karlberg L, Corell RW, Fabry VJ, Hansen J, Walker B, Liverman D, Richardson K, Crutzen P, Foley JA (2009) A safe operating space for humanity. Nature 461(7263):472–475
Rodriguez JP, Beard TD, Bennett EM, Cumming GS, Cork SJ, Agard J, Dobson AP, Peterson GD (2006) Trade-offs across space, time, and ecosystem services. Ecol Soc 11(1):28
Rodwell LD, Barbier EB, Roberts CM, McClanahan TR (2002) A model of tropical marine reserve-fishery linkages. Nat Resour Model 15(4):453–486
Rolfe J, Bennett J, Louviere J (2000) Choice modelling and its potential application to tropical rainforest preservation. Ecol Econ 35:289–302
Rosado M, Cunha-e-Sa MA, Ducla-Soares MM, Rainho Catela Nunes LM (2000) Combining averting behavior and contingent valuation data: an application to drinking water treatment (2000) FEUNL working paper no. 392. Available at SSRN: http://ssrn.com/abstract=880458
Rosenberger RS, Peterson GL, Clarke A, Brown TC (2003) Measuring dispositions for lexicographic preferences of environmental goods: integrating economics, psychology and ethics. Ecol Econ 44(1):63–76
Schenkler W, Hanemann WM, Fisher A (2005) Will US agriculture really benefit from global warming? Accounting for irrigation in the hedonic approach. Am Econ Rev 95:395–406
Seo N, Mendelsohn R, Dinar A, Hassan R, Kurukulasuriya P (2009) A ricardian analysis of the distribution of climate change impacts on agriculture across agro-ecological zones in Africa. Environ Resour Econ 43(3):313–332
Smith MD (2007) Generating value in habitat-dependent fisheries: the importance of fishery management institutions. Land Econ 83:59–73
Spash CC (2008) Deliberative monetary valuation and the evidence for a new value theory. Land Econ 84:469–488
Stern NH (2007) The economics of climate change: the Stern review. Cambridge University Press, Cambridge
Sterner T, Persson UM (2008) An even sterner review: introducing relative prices into the discounting debate. Rev Environ Econ Policy 2(1):61–76
Strahan D (2009) Extreme oil: scraping the bottom of Earth's barrel. New Sci 204(2737):34–39
Sukhdev P (2008) The economics of ecosystems & biodiversity: an interim report. European Communities, Brussels
Sumaila UR (2002) Marine protected area performance in a model of a fishery. Nat Resour Model 15(4):439–451
TEEB (The Economics of Ecosystems and Biodiversity) (2009) The economics of ecosystems and biodiversity for national and international policy makers. www.teebweb.org
Tol RSJ (2005) The marginal damage costs of carbon dioxide emissions: an assessment of the uncertainties. Energy Policy 33(16):2064–2074
Treasury HM (2003) The green book: appraisal and evaluation in central government. The Stationery Office, London. Available at http://greenbook.treasury.gov.uk/
Troy A, Wilson MA (2006) Mapping ecosystem services: practical challenges and opportunities in linking GIS and value transfer. Ecol Econ 60:435–449. doi:10.1016/j.ecolecon.2006.04.007
Tschirhart J (2009) Integrated ecological-economic models. Ann Rev Resour Econ 1:381–407
Turner RK (1999) The place of economic values in environmental valuation. In: Bateman IJ, Willis KG (eds) Valuing environmental preferences. Oxford University Press, Oxford

Turner RK, Pearce DW (1993) Sustainable economic development: economic and ethical principles. In: Barbier E (ed) Economics and ecology: new frontiers and sustainable development. Chapman and Hall, London, pp 177–194
Turner RK, Pearce DW, Bateman IJ (1994) Environmental economics: an elementary introduction. Harvester Wheatsheaf, Hemel Hempstead
Turner RK, Morse-Jones S, Fisher B (2010) Ecosystem valuation: a sequential decision support system and quality assessment issues. Ann N Y Acad Sci 1185:79–101
Turner RK, Paavola J, Cooper P, Farber S, Jessamy V, Georgiou S (2003) Valuing nature: lessons learned and future research directions. Ecol Econ 46:492–510
UKCIP (2001) Socio-economic scenarios for climate change impact assessment: a guide to their use in the UK climate impacts programme. United Kingdom Climate Impacts Programme, Oxford
United Nations (2003) The handbook of national accounting: integrated environmental and economic accounting 2003. United Nations, New York
Vatn A, Bromley DW (1994) Choices without prices without apologies. J Environ Econ Manage 26:129–148
Wallace KJ (2007) Classification of ecosystem services: problems and solutions. Biol Cons 139:235–246
Walker B, Pearson L, Harris M, Mäler K-G, Li C-Z, Biggs R, Baynes T (2010) Incorporating resilience in the assessment of inclusive wealth: an example from south east australia. Environ Resour Econ 45:183–202. doi:10.1007/s10640-009-9311-7
Weitzman M (2007) A review of the stern review on the economics of climate change. J Econ Lit 45:703–724
Welsch H, Kuhling J (2008) Using happiness data for environmental valuation: issues and applications. J Econ Surv 23:385–406
Wilson MA, Howarth RB (2002) Discourse-based valuation of ecosystem services: establishing fair outcomes through group deliberation. Ecol Econ 41:431–443
Yachi S, Loreau M (1999) Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proc Natl Acad Sci 96:1463–1468
Zandersen M, Termansen M, Jensen FS (2007) Testing benefits transfer of forest recreation values over a twenty-year time horizon. Land Econ 83(3):412–440

View publication stats

ANALISI ECONOMICA DEL DIRITTO E AMBIENTE

*Enzo Di Giulio**

Studi & Ricerche

1-97

* *Scuola Superiore Enrico Mattei, Eni*

Studi, Ricerche & Documenti

Indice

Premessa **2**

1. Un’introduzione all’analisi economica del diritto **4**

1.1. Cos'è l'analisi economica del diritto.

1.2. L'area d'indagine.

1.3. Gli elementi del gioco.

2. La contrattazione **19**

2.1. Linee generali.

2.2. Gli effetti ricchezza.

2.3. I costi di transazione.

2.4. La protezione dei diritti.

2.5. L'informazione.

3. I regimi di responsabilità **34**

3.1. Linee generali.

3.2. La responsabilità per colpa.

3.3. La responsabilità oggettiva rispetto alla responsabilità per colpa.

3.4. Confronto con altri strumenti di politica ambientale.

3.5. I regimi di responsabilità e l'ambiente.

3.6. Tendenze recenti.

4. L'assicurazione e il processo **69**

4.1. L’assicurazione.

4.2. L'assicurazione e l'ambiente.

4.3. Il processo.

5. Conclusioni **85**

Bibliografia **89**

Premessa.

Le pagine che seguono costituiscono una panoramica di alcuni aspetti dell'analisi economica del diritto rilevanti per lo studio dell'ambiente. Il lavoro raccoglie, sistematizza - e, in alcuni casi, amplia e approfondisce - lezioni svolte, negli anni, nell'ambito del corso di "Economia e Politica dell'Ambiente" del Master in Management ed Economia dell'Energia e dell'Ambiente (MEDEA) presso la Scuola Superiore Enrico Mattei, Eni. Tali lezioni sono ora sintesi ed elaborazione di contributi di studiosi di analisi economica del diritto, ora riflessione personale, ora l'una cosa e l'altra insieme.

L'obiettivo è di introdurre il lettore all'analisi economica del diritto applicata all'ambiente, evidenziando il ruolo che gli elementi istituzionali possono svolgere nel contesto delle attività economiche e dei problemi ambientali. Particolare attenzione è dedicata ai problemi della contrattazione tra agenti successiva all'insorgere di un'esternalità negativa (danno ambientale), alla responsabilità civile (oggettiva e per colpa), all'assicurazione del danno ambientale. Le interrelazioni tra agenti, conseguenti a danni ambientali, vengono esaminate attraverso un "percorso" ai cui estremi vi sono, da una parte, la contrattazione privata e l'accordo, dall'altra, il ricorso ad una Corte di Giustizia ed il giudizio dell'autorità pubblica. Il fallimento della contrattazione ed il conseguente ricorso al giudice - ovvero lo spostamento da un estremo all'altro - vengono spiegati attraverso l'azione di diverse variabili: i costi di transazione, l'informazione, le modalità di protezione dei diritti, le spese giudiziali, ecc. Anche la possibilità di comportamento strategico da parte degli agenti è presa in esame, attraverso la considerazione di alcuni casi più significativi. Infine, sullo sfondo di tale quadro teorico di riferimento, si sono poste in risalto alcune recenti, concrete esperienze di protezione dell'ambiente, in Europa e negli Stati Uniti.

Il lavoro si sforza di sottolineare costantemente il ruolo del contesto istituzionale che, in vari modi, esercita la sua azione sulle relazioni economiche degli agenti. Forme alternative di protezione dei diritti, possibilità opposte di detenzione dei diritti di proprietà, regimi diversi di responsabilità civile, regole alternative di ripartizione delle spese giudiziali costituiscono una cornice giuridico-istituzionale che vincola il comportamento, assunto razionale, degli agenti.

Ho cercato di esporre i diversi temi con semplicità e ho consegnato, così, l'esposizione alla parola e, qualche volta, al grafico o alla formula. Qui desidero ringraziare gli allievi del MEDEA per il contributo che hanno dato, con i loro commenti, al miglioramento del lavoro. Naturalmente, ogni responsabilità per eventuali errori ed incongruenze è mia.

Gennaio 1997.

1. UN'INTRODUZIONE
ALL'ANALISI ECONOMICA DEL DIRITTO

1.1. Cos'è l'analisi economica del diritto.

In un giorno di sole, un ragazzino fa una passeggiata nei pressi di una vecchia miniera abbandonata. Le piogge primaverili hanno riversato nell'invaso che un tempo ospitava la miniera una notevole quantità d'acqua, sicché ora esso somiglia ad un piccolo lago. Poiché è molto caldo, il ragazzo si tuffa per una nuotata. Nel punto in cui egli si è tuffato, poco al di sotto della superficie dell'acqua, non visibile dall'esterno, vi è una sporgenza rocciosa. Il corpo del ragazzo vi batte contro con forza: ne deriva un infortunio grave. La famiglia del ragazzo denuncia la società che aveva gestito la miniera nel passato.

Questo incidente si è verificato negli Stati Uniti nel 1969 ed ha dato luogo alla causa *Hendricks v. Peabody Coal Co*. La Corte di Giustizia ha condannato la società Peabody Coal Co., che gestiva la miniera, osservando che: "The entire body of water could have been closed off with a steel fence for between $12000 and $14000. The cost was slight compared to the risk of the children involved."[1]

Questo caso naive ci introduce all'analisi economica del diritto. Riconsideriamolo e cerchiamo di afferrarne i legami con l'economia. C'è un agente che è stato danneggiato fisicamente da un incidente. In termini più generali, il benessere della collettività è diminuito. La Corte di Giustizia, allora, effettua un confronto. Da una parte considera il costo (C) che sarebbe stato necessario sostenere per evitare l'incidente (circa 12000 - 14000 US$), dall'altra il danno atteso (D), ossia l'entità del danno fisico moltiplicata per la probabilità che il danno si sarebbe verificato. Poiché la Corte stima che D è molto alto, e maggiore di C, essa sentenzia che la Società Peabody Coal Co. è responsabile.

[1] Posner, R. A. 1992, p. 167.

Ciò che vediamo è come sia possibile ricondurre un evento della vita reale, non tipicamente economico, ad un problema economico e come un evento che appartiene al contesto del diritto possa essere esaminato in termini economici. Nel caso specifico, è la Corte stessa a prendere in considerazione in modo esplicito la variabile costo e a utilizzarla in un confronto di carattere economico. In definitiva essa, dovendo giudicare se vi è una responsabilità della parte chiamata in giudizio - la Società che aveva gestito la miniera -, si pone alcune domande di natura economica: "quanto sarebbe costato prevenire l'incidente? Che entità ha il costo della prevenzione confrontato con quello del danno?" Soprattutto la Corte si interroga intorno ad un punto: "se vi fosse stata prevenzione da parte della Peabody Coal Co., il benessere della collettività sarebbe aumentato o diminuito?" Poiché, in questo caso, la prevenzione avrebbe comportato un miglioramento del benessere collettivo, la Peabody Coal Co. è condannata: il verdetto del giudice poggia su considerazioni economiche.

Possiamo a questo punto proporre la seguente definizione: l'analisi economica del diritto è lo studio dei fenomeni legali dal punto di vista dell'economia. Essa impiega l'economia come strumento di analisi, spiegazione ed interpretazione di un insieme di questioni che tradizionalmente sono di dominio del diritto. Infatti, nei fenomeni legali - ovvero in eventi e situazioni che danno luogo a processi, regolamentazioni, multe, etc. - è spesso possibile individuare dei contenuti economici elementari e, attraverso essi, reinterpretare i fenomeni stessi.

In particolare, l'analisi economica del diritto mutua dall'economia due elementi:

- le variabili, ovvero concetti quali costo, beneficio, costo di transazione, benessere sociale, utilità, etc.;

- l'assunzione di razionalità.

L'impiego di variabili economiche come strumento di interpretazione di un fenomeno consente di scomporre un fatto complesso in pochi punti focali e di porre in luce, per così dire, il telaio che lo sottende. Questo significa che l'uso di variabili economiche rende possibile orientarsi in contesti complessi, dare ad essi una forma, averne una visione chiara: comprenderli attraverso un'interpretazione. Poi c'è l'assunzione di razionalità: gli agenti sono concepiti come massimizzatori razionali del loro interesse, monetario e non. Ciò significa che vi è un'ipotesi di comportamento che, date certe condizioni iniziali, consente di fare previsioni sulle azioni degli agenti e, dunque, di influenzarle.

Un'esemplificazione intuitiva può aiutarci a vedere meglio questo punto. In un paese si registra un sensibile incremento degli incidenti stradali, in particolare in prossimità degli incroci: un numero crescente di automobilisti non rispetta il semaforo. Pertanto le autorità, ponendosi come obiettivo una maggiore disciplina della circolazione stradale, triplicano l'entità delle multe comminate a quegli automobilisti che non rispettano il semaforo. Possiamo vedere come la scelta delle autorità si fondi sui seguenti passaggi logici:

a) un automobilista che passa con il rosso ha un costo atteso - MP - pari all'entità della multa (M) moltiplicata per la probabilità di essere sorpreso da un vigile (P);

b) il beneficio - risparmio di tempo - che l'automobilista i ricava non rispettando il semaforo è stimato pari a B_i;

c) se $B_i > MP$, l'automobilista i non rispetta il semaforo (razionalità dell'agente: massimizza il suo interesse);

d) pertanto, un incremento dell'entità della multa da M a 3M, se $P > 0$, farà presumibilmente aumentare il numero di automobilisti che rispettano il semaforo.

Questo semplice esempio ci mostra non solo, di nuovo, la possibilità di scomporre un fenomeno nelle variabili economiche elementari, ma anche come la decisione che le autorità prendono per ridurre il numero di incidenti si basi sull'ipotesi di razionalità c). Alterando una delle condizioni iniziali - entità della multa - le autorità ritengono che potranno modificare il comportamento degli agenti poiché assumono che gli agenti perseguiranno il proprio interesse, con razionalità.

Vediamo così emergere un'ulteriore dimensione dell'analisi economica del diritto. Poiché si assume la razionalità delle azioni, la legge può essere concepita come un sistema di incentivi in grado di influenzare i comportamenti degli agenti: "The primary function of law, in an economic perspective, is to alter the incentives"[2]. Pertanto, le norme, le regole di diritto, le leggi, svolgono nell'analisi economica del diritto lo stesso ruolo svolto in economia dal sistema dei prezzi.

1.2. L'area di indagine.

"Economic thought about law is old, but the economic analysis of law, which relies on formal models, is new... Like the rabbit in Australia, the economic analysis of law found a vacant niche in the intellectual ecology, and filled it rapidly"[3]. L'analisi economica del diritto è, dunque, una disciplina giovane. Tradizionalmente si fa risalire la sua data di nascita al 1960, anno in cui venne pubblicato il noto articolo di Ronald Coase sulle esternalità "The problem of the social cost". Gli sviluppi successivi approfondiranno il tema delle effetti esterni ed altri, vecchi e nuovi, ne esploreranno. Oggi, molti fenomeni che sono di interesse dell'economia, ed in particolare dell'economia pubblica, vengono investigati specularmente dall'analisi economica del diritto: public utility, common carrier, mercati finanziari, diritti di proprietà, incidenti industriali ed ambientali, fallimenti del mercato, antitrust, frequenze radiotelevisive, responsabilità dei produttori, contratti e assicurazioni sono alcuni dei problemi studiati. Tale

[2] Posner, R. A. 1992, p. 265.
[3] Cooter, R. D. and Rubenfield, D. L. 1989.

ampiezza nel range dei temi trattati dipende, essenzialmente, dalla pervasività delle regole, della legge e della politica pubblica. Più in generale, le azioni degli agenti non appartengono interamente al dominio del privato ma si svolgono nell'ambito di contesti istituzionali che le orientano e le condizionano. Tale ruolo di orientamento delle azioni private viene svolto dal contesto istituzionale non solo dinamicamente - ovvero nell'ambito di situazioni nelle quali il cambiamento di una norma o di una regola genera mutamenti dei comportamenti privati (come si è visto nell'esempio dell'automobilista e della multa) - ma anche nell'ambito di situazioni statiche. In altre parole, non occorre la variazione affinché il contesto istituzionale operi la sua azione, piuttosto esso costituisce una sorta di sfondo, una cornice nella quale le azioni private degli agenti si svolgono. Ora, poiché esiste il contesto istituzionale e dunque i comportamenti privati non avvengono in astratto, esiste la possibilità di interpretare un gran numero di fenomeni sociali studiando il ruolo che la cornice istituzionale svolge: di qui l'ampiezza dei temi che possono essere analizzati dall'analisi economica del diritto.

L'esistenza di uno standard ambientale che vieta ad un'impresa di emettere nell'atmosfera più di una certa quantità di sostanze inquinanti costituisce un esempio di influenza diretta esercitata dal contesto istituzionale. L'impresa ha di fronte una regola che stabilisce ciò che è lecito e ciò che non lo è, e con tale regola deve confrontarsi. D'altra parte, anche quando non esiste una regola diretta ed esplicita che vieta un comportamento, esiste sempre la possibilità di porla in essere. Ad esempio, i cittadini di un comune nel quale opera un'azienda ad attività fortemente inquinante possono coalizzarsi, porre il problema alle autorità pubbliche e, qualora le loro proteste venissero accettate, imporre all'azienda un differente comportamento (ad esempio, l'adozione di un filtro contro le emissioni inquinanti). In senso lato, quindi, il contesto istituzionale esercita sempre la sua azione, essendo un'entità modificabile e, entro certi limiti, mobile: le autorità pubbliche, lo stato, costituiscono una sorta di "sovragente" - idealmente super partes - al quale è possibile rivolgersi non solo per far rispettare norme già in vigore o per dirimere contese private intorno a casi incerti, ma anche per modificare regole già esistenti o per introdurne nuove. Pertanto, il contesto istituzionale è qualcosa di più dell'insieme delle regole esistenti includendo esso anche la possibilità che le regole del gioco vengano modificate.

In termini generali, possiamo dire che un contesto istituzionale è costituito da due elementi:

- sistema legale: l'insieme di procedure e di regole di funzionamento che consente all'ordine sociale di esistere e di sopravvivere (promulgazione ed enforcement delle leggi, ovvero il modo in cui le norme vengono generate e fatte osservare).

- relazioni legali: l'insieme di relazioni tra agenti riconosciute dalla società.

E' chiaro che la stabilità e la solidità delle relazioni legali tra i componenti di una certa collettività dipende crucialmente dalla stabilità del sistema legale, ovvero dell'ordine sociale. Vediamo meglio questo punto considerando un caso concreto. L'agente A è detentore di un diritto, ad esempio è proprietario di un pezzo di terreno agricolo. Ciò significa che tra A e gli altri agenti esiste una relazione legale riconosciuta dalla società tale che, qualora necessario, A può rivolgersi all'autorità pubblica (lo stato) e chiedere che il suo diritto sia rispettato e che gli altri agenti non ledano i benefici che egli può ricavare dal possesso del suo diritto. Solo se esiste un meccanismo sociale affidabile (sistema legale) che vincola gli altri agenti al rispetto del diritto di A, tale diritto (la relazione legale) è effettivamente protetto: "Rights are not relationships between me and an object, but are rather relationships between me and others with respect to that object. Rights can only exist when there is a social mechanism that gives duties and binds individuals to those duties"[4].

Ora, posto che possono essere oggetto di studio di law and economics tutte quelle situazioni nelle quali sono presenti delle regole e posto che tali regole consistono in relazioni legali tra agenti, esaminiamo più da vicino questo aspetto e vediamo qual è il campo di indagine dell'analisi economica del diritto. Per far questo dovremo definire il campo delle relazioni legali esistenti, nel presente o nel futuro, tra gli agenti, ovvero trovare uno schema di tutte le relazioni legali possibili. Interpretiamo - come già detto - una relazione legale come relazione

[4] Bromley, D. W. 1991, p. 15.

simmetrica, ovvero come diritti di un agente correlati a doveri di altri agenti. Dati due agenti A e B, saranno quindi possibili le seguenti relazioni[5]:

- 1A. Diritto di A: aspettativa o certezza di A che B si comporterà in un certo modo nei suoi confronti.

Ad esso corrisponde:

- 1B. Dovere di B: obbligo per B di comportarsi in un certo modo nei confronti di A.

Poi avremo:

- 2A. Privilegio di A: A è libero di comportarsi in un certo modo nei confronti di B.

Ad esso corrisponde:

- 2B. Nessun Diritto di B: B non può ricorrere in alcun modo se A si comporta in un certo modo nei suoi confronti.

Il caso 1A è, ad esempio, quello di un agente che è in possesso di un appartamento e che si attende che un altro agente (B) non violerà abusivamente la sua proprietà. In modo correlato, il caso 1B contempla l'obbligo che B ha di non introdursi abusivamente nella proprietà di A.
Il caso 2A, invece, sintetizza tutte quelle situazioni nelle quali la legge, pur non attribuendo ad un agente A un diritto esplicito, gli consente una libertà di comportamento rispetto alla quale un altro agente non può appellarsi allo stato per farla cessare (2B: nessun diritto). Ad esempio, gli automobilisti che inquinando l'aria danneggiano la facciata di un edificio posto su una strada molto trafficata sono detentori di un privilegio rispetto al quale i proprietari dell'edificio nulla possono, non avendo alcun diritto.

In entrambi i casi (1 e 2) ci troviamo di fronte a situazioni statiche, ossia a relazioni legali già esistenti tra gli agenti. Ora, consideriamo l'aspetto dinamico. Avremo:

[5] La presente classificazione delle relazioni legali si basa sulla sistematizzazione di Bromley (Bromley, D. W. 1991, capitolo 2) che sviluppa concetti individuati all'inizio del secolo dallo studioso di diritto W. N. Holfeld e successivamente rivisti da J. R. Commons.

- 3A. Potere di A: A può volontariamente creare una nuova relazione legale che coinvolge B.

Ad esso corrisponde:

- 3B. Responsabilità di B: B è soggetto alla nuova relazione legale creata da A.

Poi avremo:

- 4A. Immunità di A: A non è soggetto ad una nuova relazione legale che B desidera porre in essere.

Ad esso corrisponde:

- 4B. Nessun Potere di B: B non può volontariamente creare una nuova relazione legale che coinvolge A.

I casi 3 e 4 sono analoghi ai casi 1 e 2 solo che in essi si considerano non relazioni legali già esistenti tra gli agenti bensì la possibilità che tra essi potranno esservi o meno, nel futuro, nuove relazioni legali.

In sintesi, vigono, tra le relazioni legali, le corrispondenze esposte nella tavola alla pagina seguente:

Agente A	Ad esso corrisponde	**Agente B**
1A: Diritto	"	1B: Dovere
2A: Privilegio	"	2B: Nessun Diritto
3A: Potere	"	3B: Responsabilità
4A: Immunità	"	4A: Nessun Potere

Questo quadro ci consente di fare due distinzioni: i) i numeri 1 e 2 definiscono "l'area del presente" mentre i numeri 3 e 4 definiscono "l'area del futuro", ossia relazioni legali possibili ma non ancora poste in essere; ii) i numeri dispari definiscono l'area soggetta alla regolamentazione dello stato, ovvero le "relazioni legali attive", mentre i numeri pari riflettono le "relazioni legali passive", ossia un'area in cui lo stato non è presente. Ora, le relazioni legali attive costituiscono l'area di indagine dell'analisi economica del diritto. Poiché esistono delle norme che regolano le relazioni tra singoli, o tra gruppi di agenti, vi è una base per un'interpretazione economica. Al contrario, le relazioni legali passive costituiscono una sorta di terra di nessuno nella quale le interazioni tra individui non vengono regolamentate e pertanto vi è libertà di azione. E' importante sottolineare che queste due aree non sempre equivalgono, rispettivamente, all'area dell'intervento pubblico ed al mercato. Il mercato può esistere e gli agenti privati possono gestire le loro interrelazioni indipendentemente dallo stato, attraverso contrattazioni, anche all'interno dell'area delle relazioni legali attive. In essa, infatti, i diritti degli agenti sono esplicitamente riconosciuti dallo stato ed esiste quindi un punto di partenza, una base per la contrattazione. Ad esempio, A possiede un diritto di proprietà su un terreno di accesso alla spiaggia, B ha bisogno di un passaggio sul terreno di A per andare in spiaggia: A, su richiesta di B, consente a B di passare sul suo terreno in cambio di una somma monetaria contrattata privatamente. Al contrario, possono esservi situazioni nelle quali l'esistenza di una relazione legale passiva, di un privilegio, diminuisce le possibilità di contrattazione degli agenti. Un caso tipico è quello di una risorsa ad accesso libero, esempio un lago: rispetto ad essa nessun gruppo di proprietari o di utilizzatori autorizzati è definito; la risorsa è di tutti gli agenti e ciascuno è detentore di un privilegio e, al tempo stesso, non ha nessun diritto rispetto agli altri agenti: la non attribuzione dei diritti di proprietà ad un agente, singolo o gruppo, può costituire un ostacolo ad una contrattazione tra i pescatori volta ad evitare un sovrasfruttamento della risorsa nel tempo.

Ora, definito il campo di indagine dell'analisi economica del diritto, vediamo quali elementi esso contiene e quali sono le variabili fondamentali in gioco.

1.3. Gli elementi del gioco.

I fenomeni oggetto di studio di law and economics, considerata nella sua totalità, hanno caratteristiche distinte e difficilmente possono essere ricondotti ad uno schema univoco. Restringendo il campo dei fenomeni studiati è tuttavia possibile individuare alcune variabili fondamentali ricorrenti nell'analisi e studiare il ruolo che esse svolgono. A tal proposito, lo scritto di Coase del 1960 può offrire qualche elemento di orientamento. In esso si legge: "This paper is concerned with those actions of business firms which have harmful effects on others. The standard example is that of a factory, the smoke of which have harmful effects on those occupying neighbouring properties". Viene posto, dunque, come centrale un tema: l'interazione tra gli agenti, gli effetti esterni. Il comportamento di un soggetto (A) può ridurre il benessere di un altro soggetto (B): "The real question that has to be decided is, should A be allowed to harm B or should B to be allowed to harm A? The problem is to avoid the more serious harm"[6].

Ora, minimizzare il danno significa generare - date certe condizioni iniziali - efficienza[7]. In questo contesto, il sistema legale può essere visto come un meccanismo che attraverso la definizione di diritti ex ante (relazioni legali) ed attraverso il loro enforcement - la loro protezione - crea efficienza. Più in generale, la giustizia può essere considerata un bene pubblico - quindi un bene caratterizzato da non escludibilità e non rivalità - offerto dallo stato per raggiungere un'allocazione delle risorse efficiente[8]. Naturalmente, la giustizia svolge anche altre funzioni ed è possibile interpretarla come un bene in se stesso, ovvero come un valore perseguito dallo stato per ragioni etiche[9].

---

6 Coase, R. H. 1990, p. 95 e 96.

7 In termini generali, l'efficienza concerne uno stato del mondo in cui il benessere di una collettività di agenti è massimizzato. Più rigorosamente, uno stato del mondo si dice Pareto-efficiente se, partendo da esso, è impossibile aumentare il benessere di almeno un membro della collettività senza nel contempo diminuire il benessere di qualsiasi altro membro. Analogamente, se ci riferiamo ad un cambiamento, diciamo che un passaggio da uno stato del mondo ad un altro dà luogo ad un miglioramento paretiano se il benessere di almeno un membro della collettività è aumentato mentre per nessuno vi è stata una diminuzione di benessere.

8 Fenn, P. and McGuire, A. 1994, p. 3.

9 In concreto, tuttavia, al di là di finalità quali efficienza ed etica, la giustizia - quale prodotto dell'uomo - è certamente condizionata da influenze politiche e dall'azione di gruppi di interessi. Su questo tema, una

Ma cosa significa produrre efficienza e perché lo stato è indotto a intervenire? Una volta che lo stato ha definito le relazioni legali esso potrebbe astenersi da ulteriori interventi lasciando agli agenti la libertà di eseguire privatamente scambi che massimizzano il loro benessere (come ad esempio nel caso dell'accesso alla spiaggia) o, specularmente, minimizzano un danno. Tuttavia, questo non sempre è possibile perché gli scambi tra agenti sono costosi. In senso ampio, con riferimento ad una contrattazione privata in caso di danno arrecato da un agente ad un altro, esistono costi concernenti l'acquisizione dell'informazione, il bargaining vero e proprio, il controllo del rispetto dell'accordo[10]. In linea di massima, queste classi di costi di transazione corrispondono rispettivamente al periodo precedente, contemporaneo e seguente l'accordo. Vi sono situazioni nelle quali tali costi di transazione sono così elevati da precludere la possibilità di una contrattazione. In tali casi, può essere necessario un intervento dell'autorità pubblica più esteso della semplice definizione e protezione dei diritti degli agenti. Ora, trascurando tanto le strategie di intervento fondate sugli strumenti economici (tassazione pigouviana, permessi negoziabili, ecc.) quanto quelle di command & control in senso stretto (standard e multa), consideriamo in quale altro modo le autorità pubbliche possono generare efficienza.

Per rispondere a tale domanda esaminiamo un caso concreto, articolato in misura sufficiente da introdurci al problema: la "Proposed Waste Liability Directive" della Comunità Europea (1993). Questa proposta di direttiva prevede, essenzialmente, "civil liability for damage and impairment of the environment caused by waste generated in the course of an occupational activity"[11]. Ciò significa che in tutti quei casi di significativi danni fisici, chimici o biologici prodotti all'ambiente da rifiuti, colui che ha generato i rifiuti è ritenuto responsabile del danno - indipendentemente dall'intenzione e da sue eventuali negligenze - ed è obbligato a rimborsare il

---

lettura interessante è Yeager, P. C. 1991, che investiga, anche storicamente, le distorsioni subite dalla legge nell'esperienza statunitense di regolamentazione delle attività inquinanti.

[10] Per approfondimenti sui costi di transazione, si veda il "classico" Williamson, O. E. 1987, in particolare il capitolo 1.

[11] Per approfondimenti, Brealey, M. 1993, in particolare cap. 1. Questo testo costituisce un utile punto di riferimento sulla regolamentazione ambientale in Europa. Ad eccezione del capitolo 1, sulla Comunità Europea nel complesso, ogni altro capitolo viene dedicato ad uno stato europeo e ne sintetizza la regolamentazione ambientale e le principali leggi.

danneggiato. Inoltre, è richiesto che, contro l'eventualità di tali danni e onde garantire la copertura finanziaria del rimborso del danno, il produttore dei rifiuti sia coperto finanziariamente da un'assicurazione.

Vediamo che, con riferimento all'efficienza ed alla distribuzione del benessere in una collettività, la direttiva esercita la sua azione in due modi:

- Ex ante (prevenzione): poiché il produttore di rifiuti deve compensare il danneggiato per il danno eventualmente arrecato, egli sarà indotto a prevenire danni significativi all'ambiente in tutte quelle circostanze in cui il costo della prevenzione è minore dell'entità del danno. Ciò significa che ci saranno meno danni ambientali e che, quindi, il benessere della collettività aumenterà ("effetto efficienza"). Naturalmente, colui che genera i rifiuti dovrà sostenere il costo della prevenzione mentre la collettività nel suo insieme usufruirà di un ambiente più sano ("effetto distribuzione").

- Ex post (compensazione): se si verifica il danno, colui che lo ha causato è tenuto a rimborsare i danneggiati e, quindi, a sostenere il costo dell'esternalità ("effetto distribuzione").

In altre parole, la direttiva persegue un obiettivo di efficienza - massimizzazione del benessere della collettività nel suo insieme (minimizzazione dei danni ambientali da rifiuti) - attraverso una regola che ha degli effetti anche sull'allocazione del rischio e, dunque, sulla distribuzione. Se la regola non esistesse, il danneggiante non sarebbe tenuto al pagamento del danno ed il danneggiato sopporterebbe per intero il rischio di un danno ambientale prodotto da rifiuti pericolosi. Poiché c'è la regola, il rischio viene spostato su colui che genera rifiuti, con effetti distributivi (obbligo di compensazione). Occorre osservare che tale effetto distributivo non deve essere confuso con il perseguimento di un obiettivo di equità distributiva. La regola che assegna al produttore di rifiuti la responsabilità di eventuali danni è fondata su considerazioni di

efficienza allocativa (minimizzazione del danno); certo, essa ha anche un effetto distributivo ma se ciò sia desiderabile o meno costituisce un argomento distinto, da esaminare a parte[12].

In linea di massima, l'intervento delle autorità - attraverso una regola quale ad esempio la direttiva - consente di superare l'impasse nella contrattazione causata da costi di transazione elevati introducendo ulteriori elementi nel semplice schema della transazione privata tra agenti. Con la regola, infatti, altre variabili entrano nel gioco. Ad esempio, può essere necessario il ricorso ad una Corte di Giustizia che definirà se il danno è stato realmente causato dalla persona chiamata in giudizio e qual è la sua entità; vi saranno spese giudiziarie da sostenere; occorrerà tener conto della possibilità che il danneggiante si copra finanziariamente con un assicurazione (come prevede la direttiva) e che il danneggiato, analogamente, attivi un'assicurazione sulla sua proprietà e/o sulla sua salute; e così via. In altri termini, si avrà un aumento della complessità.

Possiamo a questo punto fare un passo avanti e, concentrandoci sulle esternalità negative (danni derivanti dallo svolgimento continuo di un'attività oppure da incidenti), evidenziare in uno schema astratto quali sono le principali variabili coinvolte.
Con riferimento alla figura 1[13], avremo:

- Sistema legale, che definisce:
  diritti di proprietà degli agenti;
  regime di responsabilità.
- A = Danneggiante, o Injurer; B = Danneggiato, o Vittima.
- Corte di Giustizia, che determina:
  responsabilità degli agenti;
  entità della compensazione.
- Costi di transazione privati (informazione + bargaining + rispetto dell'accordo).
- Spese giudiziali.
- Assicurazione del danneggiante. Assicurazione del danneggiato.

[12] Su questo punto: Polinsky, M. A. 1989, cap. 16.
[13] Schema adattato da Zweifel, P. and Tyran J.-R. 1994.

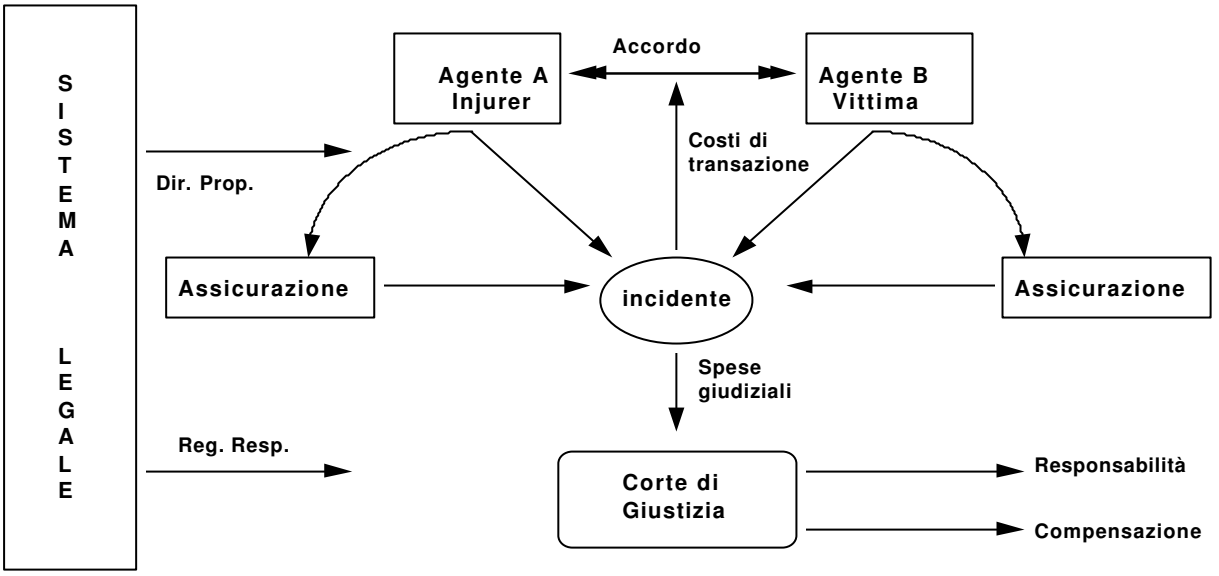


figura 1

La figura 1 prende in considerazione le relazioni tra due agenti (o gruppi di agenti) A e B nel caso che si verifichi - o possa verificarsi - un'esternalità negativa, arrecata, in tutto o in parte, da A a B. A sinistra dello schema abbiamo il sistema legale che definisce e protegge i diritti e i doveri di A e di B; nello stesso tempo, esso attribuisce poteri di giudizio alla Corte di Giustizia, stabilisce i criteri che essa deve seguire nel giudicare, definisce il regime di responsabilità; ad esempio, nel caso della proposta di direttiva europea precedentemente esaminata si stabilisce che l'injurer è da ritenersi responsabile del danno ogni volta che esso si verifica, indipendentemente dall'intenzione e dalle sue eventuali negligenze. Una volta che l'incidente si è verificato gli agenti A e B hanno due possibilità: o accordarsi circa un'eventuale compensazione e restare nella parte superiore della schema; oppure, nel caso di conflitto, spostarsi nella parte inferiore, porre all'attenzione delle autorità il problema, rivolgendosi ad una Corte di Giustizia che deciderà se A è responsabile ed in che misura lo è (compensazione). Naturalmente, questo ricorso alla giustizia avrà un costo e, in relazione al contesto giuridico vigente, gli agenti dovranno sostenere delle spese giudiziali. Tale costo rappresenta una sorta di corrispettivo dei costi di transazione privati. Il quadro è completato dalla possibilità che gli agenti A e B hanno

di rivolgersi ad un'assicurazione che, in cambio del pagamento di un premio, li indennizzerebbe qualora il danno si verificasse.

Nei paragrafi successivi, muoveremo dalla parte nord dello schema, nella quale i due agenti contrattano in quadro istituzionale semplificato, e ci sposteremo verso la parte sud considerando situazioni istituzionali più complesse. Sarà così possibile vedere, attraverso l'esame dei casi più significativi, quale ruolo ciascun elemento del gioco svolge.

2. LA CONTRATTAZIONE

2.1. Linee generali.

L'idea centrale dell'approccio di Coase è che la contrattazione tra agenti può essere uno strumento per generare un'allocazione delle risorse efficiente. In particolare, vi sono situazioni nelle quali dati due agenti A e B e data un'esternalità negativa da A a B, il benessere dei due agenti in aggregato - ovvero il benessere sociale - sarà massimizzato attraverso una contrattazione privata. Il risultato è indipendente dall'allocazione iniziale dei diritti di proprietà. Consideriamo, ad esempio, il caso di un'azienda A che inquina il campo di un agricoltore, B.

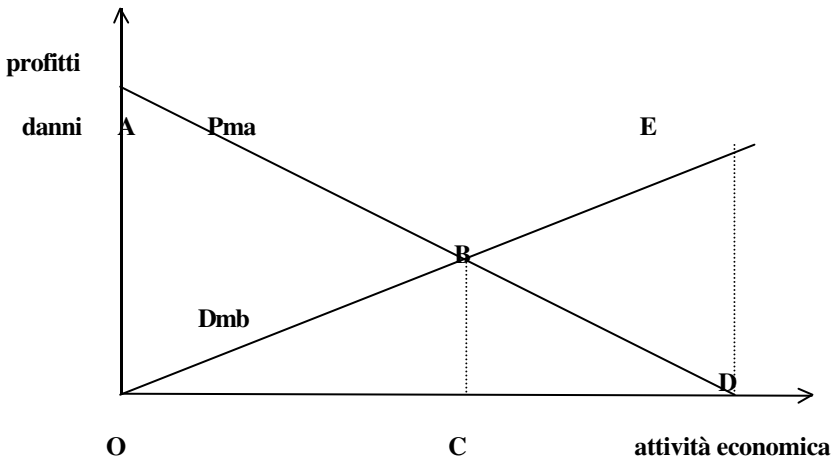


figura 2a

Sia:

Pma = Profitti marginali di A. Dmb = Danno marginale arrecato da A a B.

Nella figura 2a, il benessere aggregato di A e di B (profitto totale - danno totale) è massimo in corrispondenza di C ed è pari all'area OAB. Infatti, per ogni altro livello di produzione di A (punti a destra o a sinistra di C), il benessere sociale diminuisce. In che modo la contrattazione, consente di raggiungere il punto C e, dunque, riesce a generare efficienza?

- ***Caso 1***.

Supponiamo che A detenga i diritti di proprietà, ovvero il diritto di esercitare la sua attività produttiva senza alcun obbligo nei confronti di B. Pertanto, egli spingerà la produzione fino al punto D: massimizzerà il suo profitto totale e arrecherà una perdita OED all'agente B. Per B, sarà conveniente (o indifferente) pagare ad A una somma monetaria minore (o uguale) a CBED inducendo A a spostarsi fino a C e, così facendo, ridurrà il suo danno all'area OBC. Per A, sarà conveniente accettare l'offerta di B purché la compensazione sia maggiore o uguale alla perdita di profitto che subirebbe riducendo la produzione (BCD). Poiché per B non è conveniente offrire ad A come compensazione più di quanto egli, B, avrebbe come danno dall'attività produttiva di A, e poiché per A non è conveniente accettare come compensazione da B meno di quanto egli, A, otterrebbe come profitto dalla propria produzione, C rappresenta il punto in cui la contrattazione si arresta, generando una situazione di efficienza (benessere aggregato massimizzato). In altri termini, Dmb rappresenta la Massima Disponibilità a Pagare di B (Max DAPB), mentre Pma rappresenta, per A, la Minima Disponibilità ad Accettare una Compensazione da B (Min DACA): la contrattazione ha termine quando l'offerta di compensazione di B incontra la domanda di compensazione di A (Max DAPB = Min DACA).

- ***Caso 2***.

Il ragionamento è speculare: B possiede i diritti di proprietà e ha quindi il potere di vincolare A al punto O (produzione nulla). La contrattazione questa volta partirà da O. Pma e Dmb saranno, rispettivamente, la Massima Disponibilità a Pagare di A (Max DAPA) e la Minima Disponibilità, per B, ad Accettare una Compensazione da A (Min DACB): la contrattazione terminerà nel punto in cui Max DAPA = Min DACB, ovvero nel punto C. Il benessere aggregato è nuovamente massimizzato.

E' rilevante osservare, per entrambi i casi, che la somma effettivamente pagata da A a B (o viceversa) non è determinabile in modo rigido. In altre parole, la differenza tra DAP e DAC, a sinistra e a destra del punto C, fa sì che la compensazione non possa essere definita

univocamente: qualora l'accordo si realizzasse essa sarà una somma minore o uguale alla DAP e maggiore o uguale alla DAC dei due agenti (indipendentemente da chi detiene i diritti di proprietà). Ciò che gli agenti riusciranno ad ottenere dalla contrattazione dipenderà dalla loro abilità nel trattare. Ma non solo: vedremo, infatti, in seguito, che i comportamenti strategici degli agenti, tesi a formarsi una reputazione di contraenti forti, sono influenzati crucialmente dal modo in cui i diritti sono definiti e protetti, e dal grado di informazione delle autorità.

L'approccio di Coase, esemplificato nell'esempio precedente, ci consegna un messaggio e, al tempo stesso, ci pone un interrogativo. Il messaggio è che il mercato, lo scambio tra gli agenti, rappresenta uno strumento per raggiungere l'efficienza anche quando sono presenti esternalità. L'interrogativo concerne la misura: fino a che punto il mercato consente di raggiungere l'efficienza e di rimanere, pertanto, nella parte superiore dello schema della figura 1 evitando il ricorso all'intervento delle autorità? Il problema è tanto di carattere astratto quanto di carattere empirico. In astratto, occorrerà definire meglio quali condizioni devono vigere affinché il teorema di Coase sia vero; empiricamente, bisognerà controllare in che misura nel mondo reale sono presenti situazioni analoghe a quelle nelle quali il teorema di Coase funziona.

Con riferimento al primo punto, che qui ci interessa, occorrerà certo assumere che vi sia razionalità perfetta degli agenti e che i diritti di proprietà siano definiti in modo inequivoco e che, quindi, essi circoscrivano esattamente i diritti e i doveri. Tuttavia, ciò non è sufficiente perché altri elementi influenzano le relazioni tra i due agenti e, dunque, l'esito della contrattazione.

2.2. Gli effetti ricchezza.

Gli effetti ricchezza influenzano la contrattazione alla Coase. Un agente è condizionato da un effetto ricchezza quando, date due opzioni di scelta (X_1 e X_2):

a) non esiste somma di denaro definita che possa indurlo a spostarsi, compensandolo, da X_1 a X_2 (o da X_2 a X_1);
b) la somma necessaria a compensarlo per uno spostamento da un'opzione all'altra cambia se egli viene in possesso di un ammontare di ricchezza aggiuntiva;
c) egli non dispone di una somma di denaro sufficiente per finanziare uno spostamento dall'opzione meno preferita a quella più preferita.

a) prende in considerazione tutte quelle situazioni - come ad esempio, il rischio di morte o di gravi incidenti fisici - nelle quali qualsiasi compensazione monetaria è insufficiente. b) si riferisce a situazioni in cui le preferenze dell'agente sono influenzate da variazioni nella sua ricchezza (ad esempio, una vincita alla lotteria, un'eredità, ecc.). c), infine, riflette i casi in cui l'agente non dispone di risorse finanziare per finanziare la scelta dell'opzione preferita[14].

Ora, quando sono presenti tali effetti ricchezza la contrattazione alla Coase è ostacolata, essenzialmente perché essi influenzano la DAP e la DAC degli agenti. In a), non è possibile lo scambio moneta-opzione perché la DAC è infinita. In b), la DAP e/o la DAC degli agenti sono influenzate da elementi esterni alla contrattazione. In c), la DAP degli agenti rappresenta una DAP "teorica": il calcolo economico suggerisce una certa DAP ma essa non può essere finanziata. In generale, pertanto, tali effetti ricchezza ampliano il gioco della contrattazione tra i due agenti prendendo in considerazione situazioni - in molti casi realistiche - che riducono la possibilità che gli scambi tra agenti generino efficienza.

2.3. I costi di transazione.

Come già detto i costi di transazione possono rendere impossibile un accordo tra le parti. Al di là di questo caso estremo, è interessante esaminare il modo in cui i costi di transazione influenzano l'esito della contrattazione. Riconsideriamo i casi 1) e 2).

[14] Milgrom P. and Roberts J., 1.992, p. 35 - 36.

- ***Caso 1, con costi di transazione***.

A possiede i diritti di proprietà: ciò significa che B, se vorrà migliorare la propria situazione, dovrà sostenere i costi di transazione (ad esempio, recarsi da A, convincerlo a trattare, convincerlo a ridurre la propria attività produttiva, ecc.). Il caso estremo è quello in cui i costi di transazione sono così elevati da eccedere il beneficio, in termini di riduzione di danno, che B otterrebbe trattando con A. In tal caso, per B è conveniente non far nulla. E' possibile, però, che, pur dovendo sostenere un costo per la contrattazione, per B sia conveniente trattare perché tale costo è inferiore al beneficio della contrattazione. In questo caso B dovrà impegnare una parte delle proprie risorse per sostenere i costi di transazione. Ciò significa che la somma che egli è ora disposto a offrire ad A come compensazione (la sua DAP) è inferiore a quella che si avrebbe in assenza di costi di transazione. Max DAPB, perciò, sarà rappresentata da una nuova curva, sottostante quella dei danni marginali. La nuova DAP di B sarà, ad esempio, rappresentata da OE' (fig. 2b). Il punto di equilibrio è ora K e ad esso corrisponde un'attività produttiva pari a OC'.

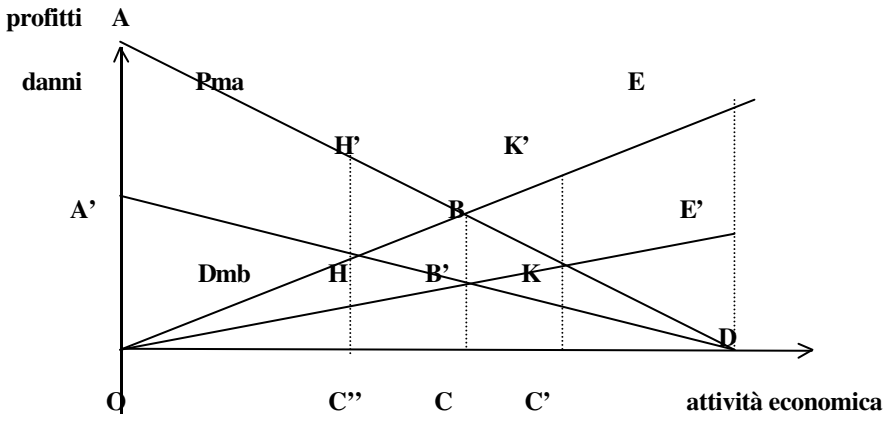


figura 2b

- ***Caso 2, con costi di transazione***.

La situazione è rovesciata. B possiede i diritti di proprietà e pertanto A dovrà sostenere i costi di transazione. Analogamente al caso 1, si ha una nuova Max DAPA (A'D), un nuovo punto di equilibrio della contrattazione (H) al quale corrisponde un nuovo livello di attività produttiva OC''.

Il caso 1 ed il caso 2, rivisti tenendo in considerazione i costi di transazione, suggeriscono due osservazioni - concernenti rispettivamente l'efficienza e la distribuzione -, oltre quella tradizionale di probabile impossibilità di contrattazione. In particolare, il contesto istituzionale, ovvero la situazione istituzionale dalla quale si parte:

- a) condiziona, attraverso i costi di transazione, l'esito della contrattazione. Non è infatti più vero che l'accordo è indipendente dall'allocazione iniziale dei diritti di proprietà. Al contrario, sono ora possibili due punti di equilibrio (OC', oppure OC");

- b) genera un effetto distributivo. Nel caso 1, B perde potere contrattuale e quindi, come conseguenza, deve sopportare un ammontare di danno maggiore di quello che avrebbe sopportato in assenza di costi di transazione (OK'C' > OBC). Nel caso 2, il contrario: la produzione di A diminuisce (OC" < OC) e ugualmente i suoi profitti (OAH'C" < OABC).

L'effetto b) evidenzia la "non neutralità" del contesto istituzionale, in particolare dello status quo, rispetto alla distribuzione del benessere[15]. L'effetto a), invece, pone un problema diverso: OC' e OC" devono essere considerati entrambi punti di efficienza? Possiamo prescindere dal contesto istituzionale e, considerandolo un semplice dato di partenza - una variabile esogena da prendere come data -, accettare come efficiente (massimizzante il benessere sociale) qualsiasi esito di contrattazione? Una risposta positiva significherebbe - salvo i casi in cui 1) e 2) dessero luogo a livelli di benessere aggregato identici (ciò dipenderà dall'andamento delle curve OE' ed A'D) - porre sullo stesso piano livelli di benessere sociale diversi. Non solo: seppure le due allocazioni iniziali dei diritti di proprietà generassero identici livelli di benessere sociale ci troveremmo in ogni caso di fronte a livelli di attività produttiva - e di corrispondente danno ambientale - diversi. Ad OC" corrisponde, comunque, un livello di inquinamento minore di OC'. Ciò pone altre domande: posto che i due livelli di benessere siano identici, dobbiamo ritenere OC" più efficiente di OC'? Estendendo la domanda: dobbiamo ritenere che

[15] Per approfondimenti sul ruolo del contesto istituzionale in economia ambientale, Bromley, D. W. 1991, in particolare capitoli 3, 4, 9.

l'allocazione dei diritti di proprietà del caso 2 sia più efficiente di quella del caso 1? E più in generale: è bene estendere la discussione sull'efficienza anche al contesto istituzionale? [16]

Naturalmente, tali questioni sollevano numerosi problemi, in certa misura coinvolgenti i valori e, dunque, non confinabili nell'ambito dell'economia intesa in senso stretto. Certamente, una considerazione da fare concerne la prospettiva nella quale considerare l'ambiente. In una prospettiva strettamente antropocentrica sarà rilevante esclusivamente il benessere dei due agenti, o perché solo gli umani contano oppure perché il loro benessere riflette ed incorpora anche il "benessere" dell'ambiente. Al contrario, in una prospettiva biocentrica il danno fisico avrà un valore di per sé e, pertanto, il benessere totale non potrà essere dedotto dalla semplice aggregazione del benessere dei singoli agenti[17] e l'ambiente dovrà essere considerato come una sorta di terzo agente da tener presente nella contrattazione. Ma qui sorgono altri problemi: come è possibile far "partecipare" l'ambiente alla contrattazione? Come valutare il danno ad esso arrecato? Non sarà sempre necessario che gli esseri umani lo rappresentino? Chi, in particolare, potrebbe rappresentarlo? E seppure tale "rappresentanza" fosse possibile, non si torna così, di nuovo, all'antropocentrismo? Qui, al di là delle risposte che ognuno può dare, occorrerà tenere presente questa possibilità: che la considerazione dei costi di transazione nell'ambito di una contrattazione alla Coase pone il problema della non-biunivocità tra efficienza economica e danno all'ambiente. Considerare il contesto istituzionale e, dunque, i diritti di proprietà come elementi esogeni del gioco può significare accettare, come efficienti, equilibri ai quali corrispondono danni all'ambiente diversi, non discriminando tra gli stessi[18].

---

[16] Il problema, in parte, somiglia alla questione della non confrontabilità dei punti sulla frontiera di Pareto. Là, la condizione iniziale sulla cui accettazione, o meno, si discute, è l'endowment iniziale degli agenti. Qui, invece, il problema viene ampliato e la discussione viene estesa al contesto istituzionale. Su questo punto, Polinsky, M. A. 1989, in particolare cap. 3 e 4, dove il problema è trattato in termini più ampi e, prescindendo dalla considerazione dell'ambiente in senso lato, il contesto istituzionale diviene esplicitamente oggetto di un giudizio di efficienza.

[17] Naturalmente, in questa trattazione semplificata tutti i problemi di passaggio dalla moneta all'utilità, e di confronto e di aggregazione del benessere di agenti diversi, vengono trascurati e si assume che la moneta sia il metro comune per misurare il benessere degli agenti. Analogamente, anche il problema della valutazione monetaria del danno subito dagli agenti non viene affrontato e l'esternalità viene sin dall'inizio espressa in termini monetari.

[18] In generale, la questione adombra il problema più ampio di cosa sia un ottimo economico. Trattare i costi di transazione come una semplice condizione iniziale condurrà a ritenere ottimali anche quelle soluzioni nelle quali un accordo non viene raggiunto: ad esempio, i costi di transazione sono elevatissimi, quindi non è conveniente contrattare, quindi non fare nulla è la soluzione ottimale. In questo modo, il teorema di Coase è sempre vero e, dunque, rappresenta una tautologia.

2.4. La protezione dei diritti.

Il contesto istituzionale esercita la sua influenza sulla contrattazione anche in un altro modo. Come evidenziato in precedenza, al sistema legale è affidato il compito di definire i diritti di proprietà e di proteggerli. Ora, il modo in cui un diritto (entitlement) è protetto svolge un'influenza rilevante nei processi di contrattazione. Più precisamente, se vi è comportamento strategico da parte degli agenti, l'esito della contrattazione dipende dal modo in cui la protezione del diritto viene garantita. Le possibilità di protezione di un diritto sono le due seguenti:

- ***Ingiunzione***: proibizione di un'azione. Se B (danneggiato) è titolare del diritto egli può proibire ad A (injurer) di causargli un danno.

- ***Compensazione del danno***. Se B è titolare del diritto, A è obbligato a compensarlo per un ammontare di denaro deciso da un'autorità pubblica (ad esempio, Corte di Giustizia); tuttavia, B non può proibire - come nell'ingiunzione - che il danno venga causato.

E' rilevante fare due osservazioni. Sia nel caso di ingiunzione che di compensazione, esiste sempre la possibilità che gli agenti contrattino e si accordino privatamente. Inoltre, l'ingiunzione e la compensazione non equivalgono necessariamente all'attribuzione di un diritto assoluto ad uno dei due agenti. Le autorità potrebbero concedere agli agenti, qualora lo reputassero necessario, dei diritti intermedi. Ad esempio, all'agente B potrebbe essere concesso un potere di ingiunzione rispetto ad A solo al di sopra di un certo livello di attività produttiva. Ciò significa che A può esercitare la sua attività produttiva liberamente fino, ad esempio, ad OC (fig. 2c) perché il potere di B di proibirgli di produrre scatta solo dopo OC. Analogamente, l'obbligo di compensazione del danno può esservi solo per un livello di danno superiore a quello corrispondente ad uno specifico livello di produzione deciso dall'autorità pubblica.

Esaminiamo, ora, alcuni casi:

- ***Caso semplice***.

In assenza di costi di transazione e di comportamento strategico da parte degli agenti, e indipendentemente da come i diritti sono protetti, si avrà una contrattazione tradizionale alla Coase, e dunque, una soluzione efficiente. Tanto per l'agente A quanto per l'agente B, indipendentemente da chi di essi sia detentore del diritto e da come lo stesso venga protetto, è conveniente raggiungere un accordo che consente ad A di produrre in corrispondenza del punto di efficienza, OC (fig. 2c). La conclusione è diversa nel caso in cui si ammetta la possibilità, realistica, di comportamento strategico. Gli agenti possono, nel tentativo di crearsi una reputazione di negoziatori "duri", rinunciare a pervenire ad un accordo grazie al quale entrambi starebbero meglio.

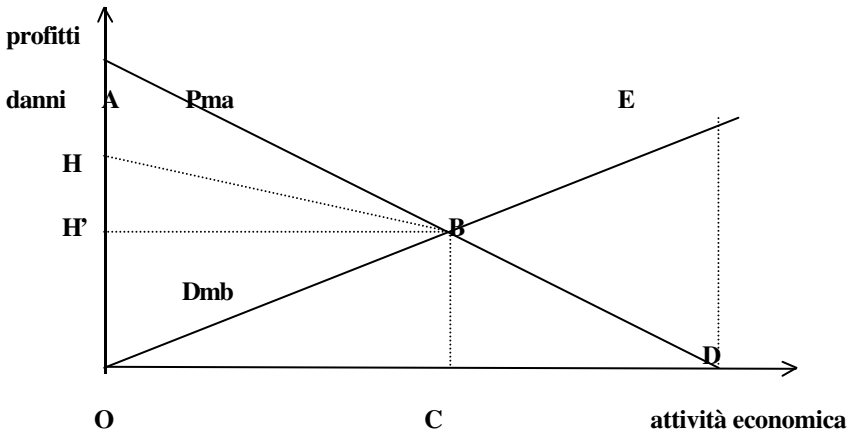


figura 2c

- ***Danneggiato protetto da ingiunzione***.

Il danneggiato potrebbe chiedere una compensazione maggiore dell'area OBC, ad esempio OHBC. Analogamente, l'injurer potrebbe rifiutarsi di compensare il danno per un ammontare maggiore di OH'BC < OHBC. La conseguenza è che il danneggiato può far applicare

l'ingiunzione, ovvero fa proibire ad A di produrre. Tale soluzione - produzione nulla e danno nullo - è Pareto-inefficiente poiché esistono altre soluzioni che avrebbero consentito ad entrambi gli agenti di disporre di un livello di benessere superiore. Se, ad esempio, il danneggiato avesse accettato da A una compensazione maggiore di OBC ma inferiore a OH'BC, vi sarebbe stato l'accordo ed entrambi gli agenti avrebbero guadagnato qualcosa. Ciò non significa che il rischio di un risultato inefficiente sia intrinseco al regime dell'ingiunzione. Infatti, l'autorità pubblica potrebbe concedere a B non un diritto assoluto ma un diritto intermedio. In altre parole, B potrebbe - anziché disporre del diritto di ingiungere ad A la totale sospensione dell'attività produttiva - essere autorizzato ad ingiungere ad A il non superamento di un certo livello di livello di produzione, comunque maggiore di zero. In tal caso, qualora le autorità facessero coincidere tale diritto intermedio con il livello di attività produttiva OC, il comportamento di B non darebbe luogo, in ogni caso, a soluzioni inefficienti. Naturalmente, tale possibilità dipende crucialmente dal grado di informazione di cui dispongono le autorità.

- ***Danneggiato protetto da compensazione***.

Quando il diritto è protetto da compensazione, la possibilità di soluzioni pareto-inefficienti dipende dall'entità della compensazione. Se infatti la compensazione richiesta ad A è pari ai danni reali, per A sarà conveniente produrre la quantità OC (fig. 2d), poiché tanto a destra quanto a sinistra di OC perderebbe qualcosa. In questo caso, il comportamento strategico di B è annullato: non c'è alcuna contrattazione da fare circa l'entità della compensazione, B non può chiedere più del danno effettivo[19] e quindi egli sarà indifferente rispetto al livello di produzione attivato da A. La situazione cambia se la compensazione viene fissata ad un livello maggiore di quello effettivo. Supponiamo che, ad esempio, alla destra di C, A deve compensare in misura pari al danno reale mentre, producendo OC, egli è tenuto a compensare B per una somma pari ad OKBC. Si avrà allora che i suoi profitti saranno ridotti a KBA

[19] Naturalmente, l'ipotesi che sia possibile determinare con esattezza ed univocità - ovvero rispetto ad entrambi gli agenti - l'entità del danno è tutt'altro che scontata. Esiste un valore univoco per un'esternalità? E, ammesso che la risposta sia positiva, come misurarlo? In altri termini, qui si prescinde da qualsiasi problema, concettuale e tecnico, di valutazione delle esternalità.

mentre l'agente B otterrà un guadagno netto (sovracompensazione = compensazione - danno reale) pari a KBO.

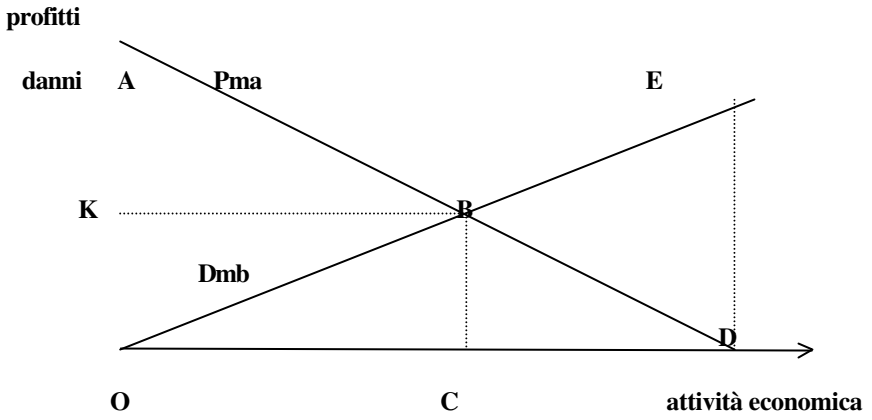


figura 2d

In questo caso, c'è spazio per il comportamento strategico: A, infatti, potrebbe negare a B la sovracompensazione, minacciandolo di non produrre. A questo punto, B avrebbe due possibilità: rendere ad A la somma "extra" KBO, oppure non rinunciare ad essa. In questo secondo caso, A - onde dare credibilità alle sue future minacce - può effettivamente decidere di non produrre negando così a B la sovracompensazione. A tal proposito è rilevante sottolineare due cose. In primo luogo, A sceglie una soluzione inefficiente (anche per se stesso) perché persegue una strategia di massimizzazione dei propri profitti di lungo periodo e, quindi, pensa che non cedere a B una volta possa rivelarsi utile per le proprie contrattazioni future: in tal senso il suo comportamento è improntato a razionalità. In secondo luogo, tanto la minaccia che A fa di smettere di produrre, quanto il comportamento di B (tenere la sovracompensazione oppure rinunciare ad essa) dipende dalle loro aspettative circa la contrattazione: poiché, come abbiamo visto, il rischio insito nel fare una minaccia è la perdita di un guadagno certo, A minaccerà B solo qualora si sentisse più forte di lui, pensando quindi di poter avere successo nella contrattazione; analogo discorso vale per B. Vale il seguente principio generale: la definizione di una compensazione pari al danno reale annulla il

comportamento strategico; al contrario, nel caso di sovracompensazione, esiste un incentivo alla minaccia per la parte ritenuta responsabile[20].

2.5. L'informazione.

Finora si è assunto che l'autorità pubblica disponga di informazione perfetta e che quindi essa riesca a fissare compensazioni uguali ai danni reali oppure a definire l'attribuzione di un diritto ad un agente in modo efficiente. In realtà ciò non sempre è vero. L'autorità, ad esempio, può non conoscere la funzione dei profitti dell'impresa o quella di danno, o entrambe. Ciò altera i risultati della contrattazione. Esaminiamo alcuni casi.

- L'autorità conosce la funzione di danno ma non conosce la funzione dei profitti. L'effetto di questa ipotesi è diverso per i due regimi.

- ***Danneggiato protetto da ingiunzione*.**

Possibilità di un risultato inefficiente. Infatti, come si è visto, affinché non vi sia comportamento strategico occorre che l'autorità definisca un diritto intermedio corrispondente al livello di produzione efficiente. Tuttavia, la non conoscenza della funzione dei profitti rende improbabile l'esatta definizione di tale diritto intermedio. Pertanto, sarà possibile che essa compia un errore e che, quindi, a causa del comportamento strategico degli agenti, l'attività produttiva si posizioni su un livello non efficiente.

[20] Per ulteriori approfondimenti, Polinsky, M. A. 1989, cap. 4, dove vengono sviluppati, attraverso esempi numerici molto semplici, alcuni casi di strategia della contrattazione. Nel presente e nel successivo paragrafo tali casi vengono riproposti, con maggiore generalità, nel continuo ed in forma grafica.

- ***Danneggiato protetto da compensazione***.

Se la compensazione è pari ai danni reali e se il diritto del danneggiato è o assoluto, o intermedio ed inferiore-uguale al livello di efficienza, l'equilibrio sarà efficiente. Infatti, se l'entitlement fosse superiore al livello di efficienza, es. OD' > OC (fig. 2e), l'impresa troverebbe conveniente produrre fino ad OD' poiché essa sarebbe tenuta a compensare il danno solo al di là di OD'. D'altra parte, B avrebbe interesse a compensare A con una somma $, tale che BCD'F ≤ $ < BCD'E', incentivandolo a produrre solo fino ad OC. Tuttavia, a causa del comportamento strategico, A e B potrebbero non accordarsi e quindi il livello di produzione potrebbe rimanere pari a OD', un risultato inefficiente. Al contrario, se l'entitlement viene fissato ad un livello minore o uguale ad OC, l'impresa troverà conveniente produrre al livello OC poiché alla destra di OC dovrà sempre compensare B per il danno reale procuratogli. Pertanto, qualora vi sia informazione imperfetta sulla funzione dei profitti, le autorità possono ugualmente generare una soluzione efficiente attraverso la semplice definizione di un entitlement assoluto.

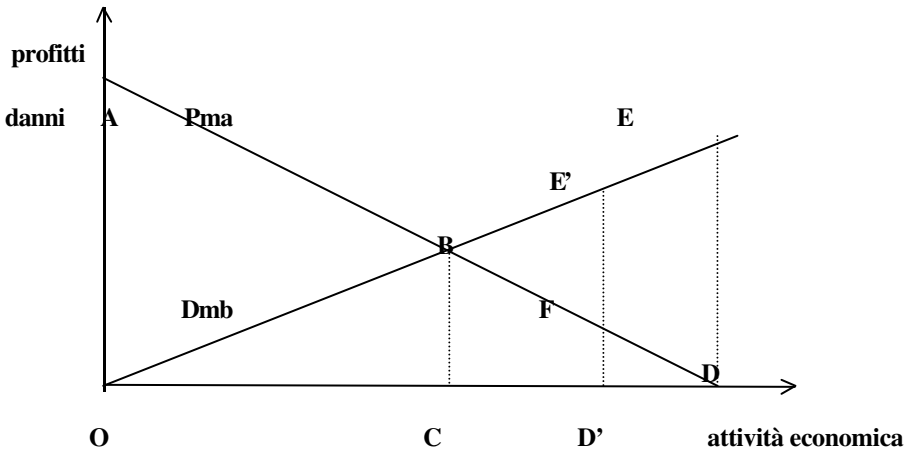


figura 2e

- L'autorità conosce la funzione dei profitti ma non quella di danno.

- ***Danneggiato protetto da ingiunzione***.

Possibilità di un risultato inefficiente. Infatti, le autorità possono non riuscire a definire un entitlement intermedio pari al livello efficiente e, quindi, emerge la possibilità di comportamenti strategici e di inefficienza.

- ***Danneggiato protetto da compensazione***.

Supponiamo che l'autorità ritenga che la funzione di danno sia più bassa di quella reale, OK anziché OE (fig. 2f).

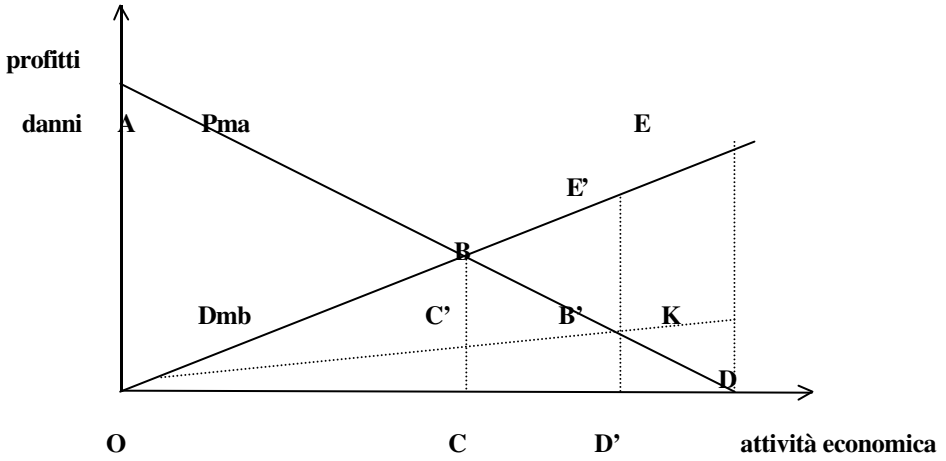


figura 2f

In tal caso, per l'azienda sarà conveniente spingere la sua produzione fino al livello OD'. D'altra parte, il danneggiato troverà conveniente indurre l'azienda, compensandola, a ridurre la propria produzione fino ad OC. Infatti, egli guadagnerebbe, non considerando il suo pagamento all'azienda, BC'B'E' (riduzione del danno - compensazione da A a B). Oppure, se il danneggiato versasse all'azienda, affinché essa riduca la propria produzione, una somma $, tale che BB'C' < $ < BC'B'E' entrambe le parti incrementerebbero il proprio benessere. Ora, la possibilità di comportamento strategico può ostacolare tale risultato inibendo il raggiungimento di OC e, dunque, lasciando il sistema su un livello produttivo inefficiente. Infatti, l'azienda potrebbe non essere disposta ad accettare una compensazione inferiore a X (essendo BB'C' < X < BC'B'E'), mentre il danneggiato potrebbe non essere disposto a compensare più di Y (essendo Y < X).

In generale, in caso di sottostima del danno non è possibile dire se uno dei due regimi si riveli, in assoluto, migliore dell'altro. L'ingiunzione può essere più efficiente della compensazione e viceversa, essendovi in entrambi i regimi possibilità di comportamento strategico e quindi incertezza nel risultato finale.

3. I REGIMI DI RESPONSABILITÀ

3.1. Linee generali.

In precedenza si è visto come il modo in cui un entitlement è protetto dalle autorità influenzi il comportamento degli agenti e, dunque, abbia un effetto in termini di efficienza. Si è visto, inoltre, come in alcuni casi un entitlement possa essere protetto da una compensazione che l'injurer versa alla vittima: uno dei due agenti, in tal caso, viene reputato responsabile di un certo evento e ciò comporta certe conseguenze sul piano monetario. Ora, nella definizione di tale responsabilità degli agenti le autorità dispongono di gradi di libertà, ovvero possono scegliere tra diversi regimi. In altre parole, l'attribuzione della responsabilità di un evento ad un agente costituisce un fatto complesso che obbliga le autorità a compiere delle scelte. Un esempio tipico è il seguente: A, automobilista, investe B, pedone. E' evidente che in questo caso l'autorità, prima di poter assegnare la responsabilità a A o a B, deve compiere una scelta in merito ad alcune questioni. Ad esempio, le seguenti: l'automobilista è sempre responsabile? Se sì, perché? Se no, in quali casi? Si può ritenere il pedone responsabile di un incidente che lo danneggia? E così via. Naturalmente, a tali domande si aggiungono quelle di carattere economico circa le implicazioni di ogni ipotetica risposta sull'efficienza. E' importante osservare che qui il discorso non si riferisce alla definizione di un insieme dettagliato di regole di comportamento che sanciscono ciò che è corretto e ciò che costituisce reato o infrazione (es. l'automobilista deve guidare a non più di 50 Km. orari nei centri abitati, deve consentire ai pedoni di attraversare le strisce pedonali, ecc.). Piuttosto, si tratta di definire schemi generali che regolino l'assegnazione della responsabilità agli agenti, in particolare in caso di danni, ambientali e non, provocati da incidenti. Ora, nella realtà giuridica dei paesi, tali schemi generali, denominati regimi di responsabilità, sono due: il regime della responsabilità per colpa (Negligence Liability) ed il regime della responsabilità oggettiva (Strict Liability). Avremo:

- ***Responsabilità per colpa***. L'agente è ritenuto responsabile dell'evento (incidente, danno all'ambiente o ad altri agenti, ecc.) se nello svolgimento della propria attività non ha preso un certo livello di precauzione (due care, o reasonable care), implicito o esplicito.

- ***Responsabilità oggettiva***. L'agente è ritenuto responsabile dell'evento (incidente, danno all'ambiente o ad altri agenti, ecc.) ogni qualvolta esso si verifica, indipendentemente dalle precauzioni che egli adotta nello svolgimento delle sue attività.

Precauzione qui significa sia adottare certe misure di prevenzione del danno (particolari tecnologie, dispositivi di sicurezza, ecc.) sia essere attenti, nel senso letterale del termine, nello svolgimento della propria attività (ad esempio, l'addetto al controllo delle pressioni di un impianto deve esercitare questo lavoro con cura evitando di distrarsi, addormentarsi, ubriacarsi, ecc.). Ora, quando vige il regime di responsabilità oggettiva il semplice verificarsi dell'incidente implica "meccanicamente" l'esistenza di una responsabilità giuridica da parte di colui che lo causa, anche se egli è stato molto attento ed ha adottato tutte le precauzioni necessarie - ma evidentemente non sufficienti - per evitarlo. Nel caso della responsabilità per colpa, invece, l'agente viene ritenuto responsabile solo se non ha preso certe misure precauzionali, ritenute ragionevoli oppure esplicitamente indicate dalle leggi. Nelle parole di Posner: "Strict liability means that someone who causes an accident is liable for the victim's damage even if the injury could not have been avoided by the exercise of due care..."[21] . In linea di massima, alcuni paesi si ispirano al regime della responsabilità per colpa mentre altri seguono il criterio della responsabilità oggettiva. Non è escluso, comunque, che i due criteri possano convivere all'interno di uno stesso sistema giuridico. Le autorità pubbliche, ad esempio, potrebbero riservare regimi diversi ad attività economiche diverse, discriminando, ad esempio, sulla base del grado di pericolosità dell'attività.

3.2. La responsabilità per colpa.

[21] Posner, R. A. 1992, p. 175.

Il principio sotteso a questo regime è quello di "nessuna responsabilità senza colpa": per dichiarare responsabile un agente occorre provarne la colpevolezza. Ciò pone alcuni problemi perché introduce, in certa misura, discrezionalità. Quando un agente è colpevole, e quindi responsabile? Cosa significa che un agente deve adottare un livello di "precauzione ragionevole" (reasonable care)? Esaminiamo un caso[22].

- Causa Blyth v. Birmingham Water Works. Alcuni giorni di gran freddo provocano la rottura dei tubi dell'acqua di una casa causando danni all'abitazione. Problema: bisogna ritenere colpevole la compagnia che gestisce la rete idrica perché non ha collocato i tubi dell'acqua abbastanza in profondità in modo tale da proteggerli dal freddo?

La Corte di Giustizia ha ritenuto non responsabile la compagnia che gestisce la rete idrica perché l'abbassamento di temperatura che ha causato il danno non aveva precedenti e, quindi, la probabilità di danno era molto basso. Ora, concentrandoci sull'aspetto economico, vediamo che queste decisioni della Corte di Giustizia poggiano, implicitamente, su un calcolo economico di efficienza. In particolare, le autorità applicano la regola del giudice Hand, una regola accettata da molti esperti e commentatori di questioni legali come la definizione legale di "precauzione ragionevole" (reasonable care). E' interessante osservare come la causa Blyth v. Birmingham Water Works preceda temporalmente la sentenza del giudice Learned Hand, come se egli, autore di una sentenza divenuta storica per gli studi di law and economics, non avesse fatto altro che esplicitare un metodo già conosciuto e applicato. Nella causa United States v. Carroll Towing Co., Hand doveva decidere chi fosse colpevole dell'affondamento di un'imbarcazione avvenuto, accidentalmente, nel corso di una manovra nella quale erano coinvolte altre imbarcazioni ed un rimorchiatore[23]. Il giudice Hand assegnò la colpa ad una delle imbarcazioni perché essa avrebbe potuto evitare l'incidente ad un costo inferiore al valore del danno. Più in generale, egli stabilì che occorreva ritenere colpevole una parte qualora il costo, che essa avrebbe dovuto sostenere per evitare l'incidente (cc), fosse stato inferiore al danno arrecato alle altre parti (d) moltiplicato per la probabilità che il danno si sarebbe

[22] Posner, R. A. 1992, p. 166

[23] Per approfondimenti, Cooter, R. 1991, p. 13 - 14.

verificato (p): cc < pd. Questa sentenza è così importante perché il giudice Hand, che peraltro espresse le sue conclusioni anche in forma matematica, non fece altro che un'operazione di ottimizzazione. Il problema può essere reimpostato nei seguenti termini. Sia:

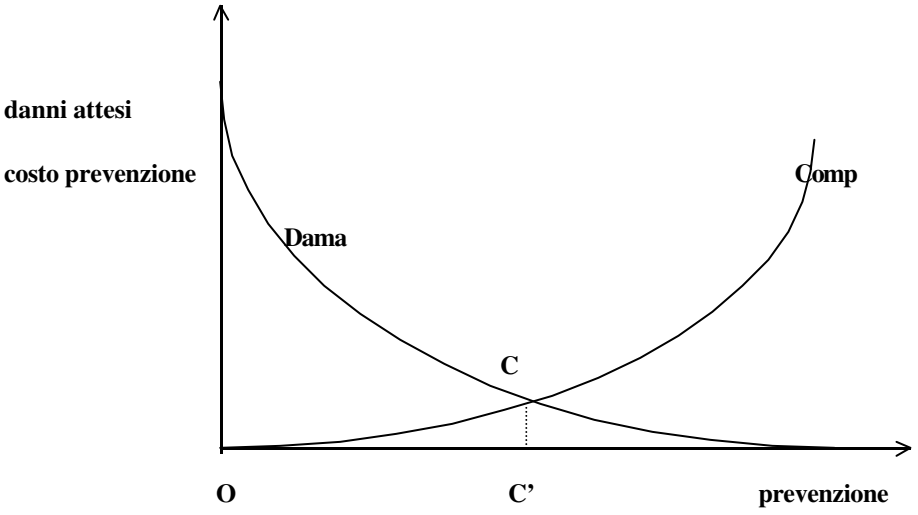
C = prevenzione.

CC = f(C), costo totale della prevenzione.

D = danno procurato dall'injurer alla vittima.

P = f(C), probabilità del danno.

Il problema è quello di trovare il livello ottimale di prevenzione che minimizza il costo totale dell'incidente (PD + CC). Risolvendo, avremo CC' = - P' D: la prevenzione è ottimale quando il suo costo è uguale al danno atteso al margine[24]. Graficamente, indicando con Comp i costi marginali di prevenzione e con Dama i danni marginali attesi - ed ipotizzando, ad esempio, che dosi successive di prevenzione siano sempre più costose e che il loro effetto preventivo sia sempre più debole -, avremo che la prevenzione ottimale è uguale ad OC' (fig. 3a), ovvero al livello corrispondente al punto in cui le due curve si incrociano. A sinistra del punto C' l'injurer è da ritenersi colpevole, e quindi responsabile, perché la prevenzione è meno costosa del danno atteso, mentre alla destra di C' egli non ha colpa perché il beneficio della prevenzione (riduzione del danno) è inferiore al costo della prevenzione.



[24] Evidentemente, il giudice Hand ragionava in termini di valori assoluti e prendeva in considerazione situazioni di scelta discreta nelle quali il potenziale danneggiante doveva scegliere tra prevenire ed evitare l'incidente, da una parte, e non prevenire ed avere l'incidente, dall'altra. Quando invece si ipotizza che dosi successive di prevenzione diano luogo a riduzioni nella probabilità del verificarsi dell'incidente [p = f(c)], allora occorre riferirsi a valori marginali ed effettuare il confronto tra costi marginali di prevenzione e danno marginale atteso. E' per questo che nella soluzione del problema di ottimizzazione sopra esposto la regola di Hand viene espressa in termini di variazioni marginali.

figura 3a

L'attribuzione della colpevolezza ad un agente, dunque, è totalmente dipendente da un giudizio di efficienza e prescinde da considerazioni di giustizia e di morale. Naturalmente, in queste norme, e più in generale nel diritto, è presente una componente etica e di valore che, tuttavia, in questo caso essa non è coinvolta. Un agente perfettamente razionale, e neutrale rispetto al rischio[25], tenderà a conformarsi sempre alla regola di Hand perché è nel suo stesso interesse farlo: preverrà il danno ogni qual volta prevenire sarà meno costoso di compensare. Ciò, naturalmente, è vero solo se l'enforcement e la compensazione sono perfetti. Ovvero, solo se la probabilità che la Corte di Giustizia dichiari responsabile un agente negligente è pari ad 1, e solo se la compensazione è non inferiore al danno atteso, l'agente avrà un incentivo ad essere diligente. Compensazioni troppo basse o sistemi giudiziari inefficienti nell'assegnare la colpa a chi è effettivamente negligente possono compromettere l'effetto incentivante della regola di Hand. Va comunque ricordato che un possibile rimedio, per questo indebolimento della deterrenza, è costituito dai danni punitivi. Quando c'è un danno punitivo l'agente è tenuto a indennizzare la vittima per un ammontare monetario uguale alla perfetta compensazione del danno più una somma ulteriore, definita danno punitivo[26].

L'effetto deterrente della responsabilità per colpa, tuttavia, non implica necessariamente che in un mondo di perfetta razionalità, perfetta informazione sul danno, perfetto enforcement, perfetta solvenza e neutralità rispetto al rischio, gli agenti siano sempre e comunque diligenti. Infatti, nel definire il livello di precauzione ragionevole (reasonable care) la Corte di Giustizia può non prendere in considerazione i costi reali che le parti in giudizio, caso per caso, avrebbero dovuto sostenere. Piuttosto, la Corte può definire la precauzione ragionevole in

---

[25] Il concetto di attitudine al rischio di un agente può essere illustrato con il seguente esempio: un agente deve scegliere se acquistare o meno un biglietto di lotteria che costa 1.000 Lire e che, se estratto, gli consentirà di vincere, 10 milioni Lire. Assumiamo che la probabilità che il biglietto venga estratto sia pari a 0,0001. Pertanto il valore atteso della lotteria è 1.000 Lit. (ovvero, 10.000.000 * 0,0001). Si tratta di una scommessa equa perchè se l'agente scommettesse spenderebbe le 1.000 Lire necessarie ad acquistare il biglietto ma avrebbe in mano un biglietto il cui valore atteso è esattamente uguale al costo del biglietto. Ora, si dice che un agente è: neutrale rispetto al rischio se è indifferente rispetto ad accettare o rifiutare una scommessa equa; avverso al rischio, se rifiuta una scommessa equa; propenso al rischio, se accetta una scommessa equa. Il comportamento dell'agente dipenderà dalla forma della sua funzione di utilità rispetto al rischio (lineare in caso di neutralità, concava in caso di avversione, convessa in caso di propensione).

[26] Per approfondimenti, Cooter R., 1991, p. 19 - 20; Posner R. A., 1992, p. 209 - 210.

base ad un valore medio dei costi. Esempio: un agente causa un incidente (danno = 200) che avrebbe potuto evitare sostenendo un costo pari a 220; la Corte, d'altra parte, ritiene che sarebbe stato possibile evitare quell'incidente sostenendo un costo pari, in media, a 180: l'agente viene ritenuto colpevole anche se, dal suo punto di vista, egli ha rispettato la regola di diligenza. In altri termini, il sistema legale può fondare il suo giudizio non sulle capacità di prevenzione che gli agenti manifestano caso per caso, bensì sul valore medio. Ciò, indubbiamente, rappresenta un risparmio per il sistema legale perché l'acquisizione di informazioni ad hoc - per ogni singolo agente di ogni singola causa - è più costosa rispetto alla definizione di un semplice standard medio.

Al fine di definire il livello di prevenzione efficiente occorre qualificare meglio la regola di Hand. Vi è incompletezza nel dire che l'injurer è responsabile ogni qual volta egli può evitare l'incidente ad un costo inferiore a quello del danno atteso e che questo garantisce l'efficienza. Infatti, ciò rappresenta solo una condizione necessaria ma non sufficiente per la definizione di un livello di prevenzione efficiente. In altri termini, occorre considerare la possibilità che la responsabilità dell'incidente ricada anche sulla vittima perché essa ha concorso al danno. Il caso degli incidenti stradali è paradigmatico. L'automobilista che investe un pedone non è sempre e comunque colpevole. Vi sono circostanze nelle quali l'accadere dell'incidente è indipendente dalla diligenza dall'automobilista: ad esempio, chi guida è attento e rispetta con esattezza ogni norma del codice stradale e tuttavia l'incidente si verifica perché il pedone - la vittima - è negligente (es. pedone ubriaco, pedone che attraversa la strada all'improvviso, non sulle strisce pedonali, ecc.). Ora, casi del genere ci inducono a considerare con maggior attenzione la regola di Hand e a verificare se, oltre l'injurer, la vittima è responsabile. Consideriamo, come esempio, la seguente possibilità: vigente un regime di responsabilità per colpa, un'azienda chimica, A, inquina le acque di un piccolo fiume. B, un agricoltore, impiega l'acqua inquinata da A per irrigare i propri campi. L'esito, per B, è la perdita di un raccolto che vale 50 milioni di Lire. Supponiamo che l'azienda possa evitare, con certezza, di inquinare il fiume istallando un dispositivo elettronico di sicurezza e di controllo delle proprie emissioni, che costa 6 milioni di lire e supponiamo che l'agricoltore può irrigare i suoi campi attingendo l'acqua ad un altro fiume, più distante del precedente dalla sua campagna, e che ciò

comporterebbe per lui un costo di 8 milioni. Ora, siamo di fronte ad un situazione nella quale: costo prevenzione dell'injurer < costo prevenzione della vittima < danno. In questa situazione chi è responsabile? L'agricoltore potrebbe dire che l'azienda è responsabile perché avrebbe potuto evitare l'incidente ad un costo inferiore a quello che egli avrebbe dovuto sostenere. D'altra parte, l'azienda potrebbe ribattere che la spesa della prevenzione per l'agricoltore è ben inferiore al danno perché egli, con una spesa di 8 milioni, avrebbe evitato un danno di 50 milioni. Controversie analoghe vengono risolte seguendo un approccio definito "contributory negligence". Secondo tale approccio la prevenzione deve essere esercitata dalla parte i cui costi di prevenzione sono più bassi, oppure - in caso di prevenzione congiunta - da entrambe le parti al minimo costo. Pertanto, nel nostro caso, l'agricoltore non è tenuto a prevenire: è sufficiente che l'azienda, la cui prevenzione è meno costosa, installi il suo dispositivo di sicurezza e ciò eviterà l'incidente. Poiché l'azienda non lo fa, essa è responsabile ed è tenuta a compensare il danno. Naturalmente, il caso qui descritto è solo uno dei possibili. Si tratta di un caso di "precauzione alternativa" perché la soluzione efficiente è che una delle parti, l'azienda, adotti certe misure preventive e questo sarà sufficiente ad evitare l'incidente, al minimo costo. Tuttavia, ciò non sempre è vero. Esistono situazioni nelle quali il costo della prevenzione è minimizzato solo quando entrambe le parti, congiuntamente, adottano misure precauzionali. Consideriamo un altro esempio: vigente un regime di responsabilità per colpa, abbiamo, da una parte, un cementificio, dall'altra, una piccola azienda tessile. In certi giorni ventosi, il cementificio danneggia con i suoi fumi i semilavorati tessili esposti all'aperto. Ciò può procurare un danno economico di alcune decine di milioni. Supponiamo che, come prevenzione, il cementificio possa alternativamente: a) adottare un filtro, molto moderno, il cui costo è 6 milioni di Lire; b) adottarne un altro, più economico e meno moderno che costa 3 milioni; c) non far niente. D'altra parte, l'azienda tessile può alternativamente: i) non far niente e continuare ad esporre all'aria i propri semilavorati per l'intero giorno; ii) esporli per non più di cinque ore al giorno; iii) esporli per non più di due ore al giorno. Questa limitazione nel tempo di esposizione dei semilavorati rappresenta per l'azienda tessile un costo: nel caso ii) 1.000.000, nel caso iii) 8.000.000. D'altra parte, a fronte di tali costi crescenti, il rischio che i prodotti vengano danneggiati dai fumi diminuisce. In sintesi:

	Costo della prevenzione di A	Costo della prevenzione di B	Costo totale
combinazione ai)	6.000.000	0	6.000.000
combinazione bii)	3.000.000	1.000.000	4.000.000
combinazione ciii)	0	8.000.000	8.000.000

Assumiamo che ognuna delle tre combinazioni riduca il danno atteso in misura identica. L'alternativa bii), pertanto, risulta la meno costosa: essa, dunque, è anche quella che le autorità reputano "efficiente"[27]. Ora, c'è da chiedersi se il regime di responsabilità per colpa, fondato sulla "contributory negligence" indurrà gli agenti A e B a scegliere rispettivamente le alternative b) e ii). Certo, tanto per A quanto per B non sarà conveniente scegliere, rispettivamente, le alternative a) e iii). E' vero che queste due alternative, se scelte congiuntamente, minimizzano il rischio di incidente. Tuttavia, poiché siamo in un regime di responsabilità per colpa, con contributory negligence, e poiché le autorità fissano il livello di prevenzione ragionevole negli standard b) e ii), gli agenti non sono tenuti - né lo trovano economico - ad attivare una prevenzione maggiore di quella richiesta. Per il cementificio, al fine di non risarcire il danno, è sufficiente rispettare lo standard di diligenza b), mentre per l'azienda tessile, al fine di essere risarcita, basta attivare l'alternativa ii). D'altra parte, livelli inferiore di precauzione sarebbero ugualmente non convenienti per entrambi gli agenti: per A, perché sarebbe tenuto a compensare un danno di valore superiore al costo della prevenzione, per B, perché rischierebbe - qualora l'incidente si verificasse - di non essere compensato. Ecco dunque che un regime di responsabilità per colpa, con contributory negligence, è in grado di generare una soluzione efficiente[28]. E' opportuno comunque ricordare che, affinché ciò sia vero, devono

[27] Qui, per semplicità, si ipotizza che le altre possibili combinazioni siano meno convenienti rispetto alle tre mostrate nella tavola. Ciò perché o esse prevengono poco o nulla il danno [cii), bi) e ci)], oppure perché, pur assumendo che riducano il danno atteso in misura pari a quella delle combinazioni esposte nella tavola, sono più costose [aii) e aiii)].

[28] In alcuni contesti istituzionali "contributory negligence" è sostituita da "comparative negligence", un approccio che, in caso di incidente e di negligenza di entrambi gli agenti, obbliga l'injurer a compensare una percentuale (es. 90%) del danno e non la sua totalità, come nel caso di contributory negligence. E' possibile dimostrare, comunque, che gli effetti di questi due approcci sono identici e che entrambi inducono negli agenti comportamenti efficienti.

essere valide le assunzioni di perfetta razionalità, perfetta informazione sul danno, perfetto enforcement, perfetta solvenza e neutralità rispetto al rischio degli agenti. Scostamenti da queste assunzioni possono implicare un risultato diverso e, conseguentemente, allontanamenti dall'efficienza.

3.3 La responsabilità oggettiva rispetto alla responsabilità per colpa.

Il regime di responsabilità oggettiva si caratterizza, rispetto a quello della responsabilità per colpa, per un maggior rigore poiché esso non fa distinzioni tra casi ed assegna - sempre, indipendentemente dal grado di diligenza - la responsabilità all'agente che ha causato il danno. L'esistenza di una relazione di causa-effetto tra lo svolgimento di una certa attività ed il verificarsi del danno determina la responsabilità dell'agente che esercita l'attività. Questo maggior rigore della responsabilità oggettiva, tuttavia, non genera - in linea di massima - un maggior livello di prevenzione rispetto al regime della responsabilità per colpa. Consideriamo le seguenti due possibilità, 1 e 2:

	Danno atteso	Costo della prevenzione
1	100	80
2	100	130

Nel caso 1, entrambi i regimi danno luogo a prevenzione: quello della responsabilità per colpa perché, in base alla regola di Hand, l'injurer sarebbe ritenuto responsabile e quindi obbligato a risarcire la vittima di 100 > 80, e ciò non è per lui conveniente; quello della responsabilità oggettiva perché l'injurer dovrebbe compensare in ogni caso la vittima di 100 > 80, e ciò non è conveniente. D'altra parte, nel caso 2, in tutti e due i regimi è conveniente non prevenire: nel caso della responsabilità per colpa perché l'injurer non è ritenuto responsabile e, quindi, per lui non è conveniente attivare la prevenzione; nel caso della responsabilità oggettiva, perché pur essendo l'injurer ritenuto comunque responsabile, per lui non è conveniente prevenire, essendo il costo della prevenzione maggiore del valore della compensazione (130 > 100). Il risultato è

identico nel caso in cui il danno atteso sia una funzione continua della prevenzione, ovvero nel caso in cui sia necessario un confronto al margine.

Occorre comunque osservare che la considerazione delle spese giudiziali e dell'incertezza circa l'esito del processo modifica le conclusioni generali sui regimi di responsabilità. Entrambi i regimi possono fallire nel generare comportamenti ottimali del potenziale danneggiante perché, da un parte, l'incertezza intrinseca nel processo fa diminuire il danno atteso da pagare, dall'altra le spese giudiziali tendono ad aumentarlo. L'effetto aggregato di questi due fattori è incerto, tuttavia è improbabile che essi si compensino esattamente in modo tale da lasciare immutato l'effetto dei regimi di responsabilità sul comportamento del potenziale injurer.

Naturalmente, la prevenzione è solo uno degli aspetti tra quelli da esaminare in un confronto tra i due regimi. Differenze esistono per quanto concerne altri aspetti. Esaminiamo qui di seguito, in termini qualitativi, i più rilevanti.

- ***Livello di attività.***

Vi sono danni la cui entità dipende non solo dall'adozione di misure preventive, essendo essi dipendenti anche dalla misura in cui l'attività viene esercitata (es. le emissioni inquinanti di un impianto chimico dipendono dai volumi di produzione). Ora, poiché la responsabilità per colpa è un criterio fondato sulla diligenza, essa, nell'assegnare la responsabilità ad un agente, prescinde dal livello di attività. In altri termini, un agente può essere diligente, e dunque non responsabile, e nello stesso tempo causare gravi danni all'ambiente, o a terzi, a ragione di un'attività, di produzione o di consumo, esercitata intensamente. La responsabilità oggettiva, invece, imputando in ogni caso la responsabilità all'injurer, prende in considerazione il livello di attività. Ciò implica che in un contesto istituzionale poggiante sulla responsabilità oggettiva i potenziali inquinanti, nel decidere del proprio agire, presteranno attenzione non solo alla variabile "diligenza" ma anche a quella "attività". Dal punto di vista del policy maker, poi, cambia l'obiettivo: nel caso della responsabilità per colpa si ambisce a imputare all'injurer solo i costi di quegli incidenti che potevano, ragionevolmente - con diligenza -, essere evitati; nel caso di quella oggettiva, invece, il policy maker intende proprio legare la responsabilità al

verificarsi tout court del danno e, quindi, l'injurer è ritenuto responsabile anche di danni associati al semplice svolgimento di un'attività.

- ***Effetti sui prezzi e sui consumi***.

Nell'ipotesi che la prevenzione efficiente attivata dal potenziale injurer non elimini totalmente il danno - come ad esempio nella figura 3a, in cui al livello di prevenzione efficiente non corrisponde un danno residuo pari a zero -, i due regimi di responsabilità danno luogo a prezzi, e quindi consumi, diversi. Assumiamo, ad esempio, perfetta efficienza e razionalità da parte tanto delle autorità quanto dei produttori. In questo caso, come abbiamo visto, sia nell'ipotesi di responsabilità oggettiva che in quella di responsabilità per colpa, il potenziale injurer attiverà un livello di prevenzione efficiente. Supponiamo, però, che al livello di prevenzione efficiente sia associato un danno residuo. Ora, mentre nel caso della responsabilità per colpa l'injurer - avendo egli rispettato lo standard di diligenza - non è tenuto a compensare la vittima per tale danno residuo, nel caso della responsabilità oggettiva l'injurer dovrà compensare in ogni caso la vittima, anche per la parte di danno residuo. Ciò significa che in un regime di responsabilità oggettiva i prezzi dei beni incorporeranno il costo del danno residuo, mentre in un regime di responsabilità per colpa ciò non accade, poiché il produttore non compensa la vittima per la parte di danno residuo. Dunque, poiché, in quest'ultimo caso, i prezzi non riflettono i costi totali del prodotto (costi privati del produttore + costi sociali), essi indurranno un livello inefficiente di consumi. In altri termini, in un regime di responsabilità per colpa, date le ipotesi qui fatte, i consumi - e dunque, il livello dell'output - saranno superiori al livello di efficienza. Naturalmente, questa differenza tra i due regimi scompare nel caso in cui i potenziali injurer siano di fronte ad una scelta discreta - da una parte non prevenire e causare danno, dall'altra prevenire ed eliminare il danno - ed il costo della prevenzione sia inferiore al danno atteso. In questo caso, infatti, la non prevenzione implica il danno e, quindi, l'obbligo di compensazione, tanto in un regime di responsabilità per colpa quanto in uno di responsabilità oggettiva.

- ***Distribuzione***.

La precedente distinzione riflette una differenza distributiva. I due regimi ripartiscono il costo dell'incidente in modo diverso. Se vi è danno residuo, e l'injurer è stato diligente, il regime di responsabilità per colpa non dà luogo ad alcun indennizzo della vittima che, pertanto, sostiene da sola il costo dell'incidente. Al contrario, la responsabilità oggettiva trasferisce questo costo sempre sull'injurer. In termini più generali, occorre dire che i due regimi corrispondono a due schemi diversi di allocazione del rischio. Nel caso della responsabilità oggettiva la vittima è sempre assicurata contro un possibile incidente, essendo essa sempre compensata. Nel caso della responsabilità per colpa, invece, non è così e, se l'injurer è stato diligente, la vittima non verrà compensata per il danno subito.

- ***Spese giudiziali***.

Le parti non necessariamente si accordano per una soluzione privata del "conflitto". In un certo numero di casi, come già evidenziato, è indispensabile un ricorso alla Corte di Giustizia e tale ricorso comporta un costo. Ora, questo costo - spese giudiziali - non è identico per i due regimi. La responsabilità oggettiva dà luogo a spese giudiziali minori rispetto alla responsabilità per colpa, essenzialmente perché le cause sono più snelle e meno complesse: la variabile "diligenza" - che, peraltro, è quella maggiormente soggetta ad interpretazione e discrezionalità - non viene presa in considerazione. D'altra canto, la maggior semplicità delle cause, nel criterio della responsabilità oggettiva, accresce i ricorsi al giudice: poiché la diligenza del danneggiante non costituisce un elemento del giudizio, aumentano le probabilità di vittoria del danneggiato nel processo e, dunque, le citazioni in giudizio ed il numero di processi. Ora, poiché al fine di definire gli effetti dei due regimi in aggregato - rispetto alla variabile spese giudiziali -, occorrerà prendere in considerazione sia il numero di incidenti sia il numero di cause, non è possibile determinare univocamente l'effetto complessivo: in un regime di responsabilità oggettiva diminuisce il costo unitario delle cause ma aumenta il loro numero. Inoltre, l'incertezza è accresciuta dal modo in cui i due diversi regimi agiscono sul numero di incidenti: la responsabilità oggettiva, attraverso la considerazione del livello di attività, potrebbe determinare una riduzione nel numero degli incidenti; d'altra parte, però, questa è solo una possibilità perché vi sono situazioni nelle quali gli incidenti sono inevitabili e, dunque, accadono

indipendentemente dall'adozione di un elevato livello di prevenzione o da un moderato volume di attività.

Le caratteristiche dei due regimi sono evidenziate, sinteticamente, nella tavola alla pagina seguente:

Un confronto sintetico tra

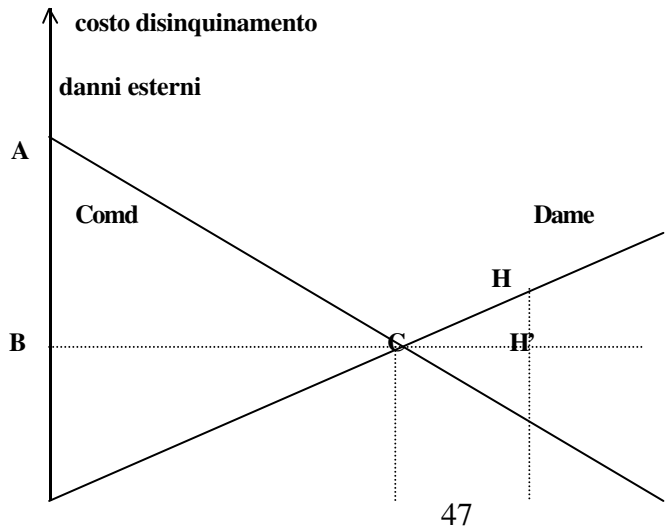
responsabilità per colpa e responsabilità oggettiva

Responsabilità per colpa.	**Responsabilità oggettiva**.
• Principio: nessuna responsabilità senza colpa.	• Principio: responsabilità indipendente dalla colpa.
• C'è colpa, e, dunque, responsabilità: quando viene violato lo standard di diligenza (implicito o esplicito).	• C'è responsabilità: quando esiste una relazione di causalità tra l'attività svolta da un agente e il danno.
• Effetti sulla prevenzione: l'agente razionale previene se il costo della prevenzione è minore del danno atteso.	• Effetti sulla prevenzione: l'agente razionale previene se il costo della prevenzione è minore del danno atteso.
• Effetti sul livello di attività: nessuno.	• Effetti sul livello di attività: l'agente, se neutrale rispetto al rischio, svolge la sua attività finché il suo profitto marginale è maggiore o uguale al danno atteso da compensare.
• Effetti sui prezzi e sui consumi: in caso di danno residuo, se l'injurer è diligente i prezzi non riflettono l'intero costo sociale e, quindi, il livello dei consumi e dell'output non è efficiente.	• Effetti sui prezzi e sui consumi: i prezzi riflettono in ogni caso l'intero costo sociale e, quindi, il livello dei consumi e dell'output è efficiente.

• Allocazione del rischio e distribuzione: l'injurer compensa la vittima solo se è ritenuto colpevole; la vittima non è assicurata contro il rischio di incidente.	• Allocazione del rischio e distribuzione: l'injurer compensa sempre la vittima; la vittima è assicurata contro il rischio di incidente.
• Spese giudiziali: "alto" costo unitario delle cause, "basso" numero di cause.	• Spese giudiziali: "basso" costo unitario delle cause, "alto" numero di cause.

3.4. Confronto con altri strumenti di politica ambientale.

Esistono alcune somiglianze tra la responsabilità per colpa e lo standard ambientale come tra la responsabilità oggettiva e la tassa ambientale. Tuttavia, mentre standard e tasse ambientali sono strumenti pensati per i casi di inquinamento continuo - ovvero per forme di danno prodotto gradualmente dallo svolgimento di una certa attività, produttiva e non-, i regimi di responsabilità prendono in considerazione, prevalentemente, fenomeni di inquinamento accidentale. E' comunque possibile, trascurando tale differenza, effettuare un confronto. Lo standard di diligenza sul quale si fonda la responsabilità per colpa può essere interpretato come uno standard ambientale tradizionale, ovvero come un limite al di là del quale comincia la responsabilità dell'injurer ed egli è tenuto al pagamento di una somma. Assumiamo anche che le autorità siano in grado di gestire in modo efficiente qualsiasi strumento di politica ambientale. Esse saranno in grado di fissare una tassa ottimale, uno standard ottimale e di gestire in modo efficiente entrambi i regimi di responsabilità.



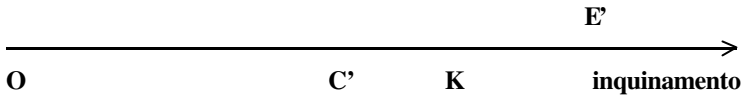


figura 3b

Con riferimento alla figura 3b, in cui le curve Comd e Dame indicano rispettivamente i costi marginali di disinquinamento e i danni marginali esterni, ciò significa:

- tassa pari a OB.
- standard in corrispondenza del livello di inquinamento OC' e multa pari a OB.
- per il regime di responsabilità per colpa, standard di diligenza pari a OC' e compensazione del danno a destra di OC'.
- per il regime di responsabilità oggettiva, compensazione integrale del danno.

Ora, un agente razionale, potenziale danneggiante:

- nel caso della tassa: i) sosterrà il costo di disinquinamento E'CC'; ii) inquinerà OC'C iii) pagherà allo stato una somma OBCC'.
- nel caso dello standard: i) sosterrà il costo di disinquinamento E'CC'; ii) inquinerà OC'C; iii) non pagherà nulla per tale inquinamento residuo OC'C .
- nel caso della responsabilità per colpa: i) sosterrà il costo di disinquinamento E'CC'; ii) inquinerà OC'C; iii) non pagherà nulla per tale inquinamento residuo OC'C .
- nel caso delle responsabilità oggettiva: i) sosterrà il costo di disinquinamento E'CC'; ii) inquinerà OC'C; iii) dovrà compensare il danneggiato per tale inquinamento residuo OC'C.

Dunque, in tale schema astratto lo standard è equivalente alla responsabilità per colpa: per entrambi gli strumenti il danneggiante è responsabile solo al di sotto di una soglia critica di disinquinamento (E'C'), corrispondente, in un mondo di perfetta razionalità e perfetta informazione, al punto di efficienza. Oltre tale soglia l'injurer, pur causando un danno, non è tenuto ad alcun pagamento. Invece, al di sotto della soglia esistono alcune differenze: mentre nel caso dello standard il danneggiante paga la multa OBH'K, in un regime di responsabilità per colpa egli è obbligato a compensare la vittima nella misura OHK. Ciò che cambia, quindi,

passando dallo standard alla responsabilità per colpa, è il destinatario del pagamento - rispettivamente lo stato e la vittima - e l'entità del pagamento (rispettivamente OBH'K e OHK).

La tassa, dall'altra parte, appare molto simile alla responsabilità oggettiva. Entrambi gli strumenti obbligano il danneggiante a pagare una somma per l'inquinamento prodotto, anche se esso è generato da un volume di attività inferiore al punto di ottimo. Di nuovo, però, vi sono differenze concernenti il destinatario e la misura. Nel caso della tassa, è lo stato che incassa - assumendo perfetta razionalità del danneggiante - OBCC'. Nel caso della responsabilità oggettiva, invece, è la vittima ad essere compensata, in misura pari a OCC'.

Occorre comunque sottolineare che tali confronti, seppur molto utili perché consentono di stabilire delle relazioni tra i diversi strumenti e di comprendere il posto che ciascuno di essi occupa, costituiscono delle semplificazioni poiché prescindono, oltreché dalla complessità del mondo reale - tanto la perfetta razionalità quanto la perfetta informazione rappresentano infatti due ipotesi molto forti -, da alcune specificità che caratterizzano i regimi di responsabilità rispetto agli standard e alla tassazione ambientale[29]. Il funzionamento dei regimi di responsabilità, in particolar modo della responsabilità per colpa è meno meccanico rispetto alla tassa ed allo standard: occorre un'azione di denuncia del danno da parte dell'agente danneggiato o di un agente autorizzato a ciò, un processo spesso complesso di valutazione dei fatti (relazione di causalità e prove di colpevolezza) da parte di una Corte di Giustizia, il coinvolgimento di avvocati, la stima dell'indennizzo. Tali fattori generano complessità e rendono i regimi di responsabilità strumenti di protezione dell'ambiente piuttosto complessi e articolati. Qui, considerata la preminenza di un obiettivo di esplorazione dell'argomento rispetto ad uno di approfondimento formale e considerato il fatto che un confronto effettuato in astratto - a ragione del fatto che ogni modello astratto non è altro che una fra le tante possibilità di interpretare la realtà - è strettamente dipende dal framework scelto, è preferibile limitarsi ad elencare alcuni punti di forza e di debolezza dei regimi di responsabilità rispetto agli strumenti. In positivo, i regimi di responsabilità:

[29] Inoltre, qui - a differenza dello schema tradizionale (fig. 3a), in cui il danno esterno viene posto in relazione con la sola prevenzione - si ipotizza che il danno esterno dipenda dall'attività economica e la prevenzione viene ridenominata disinquinamento.

- non richiedono, in ogni circostanza, informazione perfetta da parte delle autorità sulla funzione di danno: sono le vittime che si attivano per essere compensate e forniscono, escludendo casi di *cheating behaviour*, la misura del danno. Il costo dell'informazione, quindi, è spostato dal settore pubblico a quello privato: esso può essere molto elevato, di volta in volta, per gli agenti coinvolti in una contrattazione o in un giudizio, ma basso per la collettività.

- non richiedono, al contrario di sistemi di monitoraggio permanenti, che l'informazione venga raccolta in modo continuo: le Corti di Giustizia la raccolgono quando sono chiamate a giudicare, e ciò determina un risparmio sociale.

- non richiedono un meccanismo di raccolta e gestione delle compensazioni a favore delle vittime perché esse sono indennizzate direttamente dalle parti responsabili.

- non sono soggetti ad influenze di imprese *rent seeking* che, nel caso di standard e tasse, possono condizionare le scelte delle autorità.

- sono più convenienti di una regolamentazione uniforme che si rivela altamente inefficiente nei casi in cui i costi di prevenzione variano molto tra i danneggianti.

Tuttavia, vi sono degli aspetti negativi:

- si tratta di strumenti che possono essere poco efficienti nei casi di inquinamento continuo e graduale: la vittima, in questi casi, può avere maggiori difficoltà, rispetto ai casi di inquinamento accidentale, a dimostrare di essere stata danneggiata.

- quando l'onere della prova è a carico del danneggiato - ovvero è il danneggiato a dover provare che l'injurer è responsabile - vi è la possibilità che egli, a ragione di scarsa iniziativa

o di aspettative di alti costi di transazione, non ricorra in giudizio e che, quindi, l'esternalità non venga compensata.

- le Corti di Giustizia possono incontrare notevoli difficoltà nell'accertare la dinamica dell'incidente, il nesso di causalità, il grado di diligenza dell'injurer, ecc., ed i costi di gestione del sistema legale possono essere molto elevati, come rivela ad esempio l'esperienza americana recente.

- le Corti di Giustizia possono essere condizionate da *deep-pocket policy*, ovvero far dipendere l'entità del danno stimato dalla capacità di pagamento dell'accusato, così come esse la percepiscono. Secondo alcuni studi, le stime del danno sono più alte del 30% - 90% quando l'accusato è una società per azioni o un'agenzia governativa[30].

In sintesi, questi appaiono essere i principali argomenti favorevoli e contrari ai regimi di responsabilità. La "superiorità" o l' "inferiorità" dei regimi di responsabilità rispetto ad altri strumenti dipende, di volta in volta, dal modo in cui essi si manifestano e combinano. Occorre osservare, comunque, che la traduzione dello schema astratto in realtà può essere realizzata in modi diversi. Su un aspetto importante, quale ad esempio l'onere della prova, gli stati adottano soluzioni diverse. In generale, poi, il ricorso concreto allo strumento della responsabilità civile si rivela un obiettivo particolarmente complesso qualora il campo di applicazione sia costituito dall'ambiente.

3.5. I regimi di responsabilità e l'ambiente.

L'impiego dello strumento della responsabilità civile per i casi di esternalità ambientali suscita numerosi problemi essendo l'ambiente un contesto particolare, con tipicità e caratteristiche proprie. Come abbiamo visto, in uno schema nel quale, da una parte, gli agenti sono assunti perfettamente razionali, informati, solventi e neutrali rispetto al rischio e, dall'altra, vi è perfetto

[30] Per un approfondimento sui confronti tra regimi di responsabilità ed altri strumenti della politica ambientale, Zweifel, P. and Tyran, J.-R., 1994.

enforcement delle regole da parte dello stato, i regimi di responsabilità produrranno i loro effetti positivi sull'efficienza. Al contrario, scostamenti da queste assunzioni implicheranno, salvo eccezioni, fallimento o indebolimento della funzione di generazione dell'efficienza. Ora, al di là dell'irrealismo complessivo di tali assunzioni - il cui ruolo non è certo quello di costituire uno specchio dei fatti reali, bensì di consentirne la modellizzazione -, i problemi di impiego dei regimi di responsabilità nel campo dei fenomeni ambientali hanno riguardato, in larga misura, un'assunzione in particolare: quella di perfetta informazione. Certo, nel reale, anche le altre assunzioni si scostano dal modello ideale e si deformano in misura tale da rendere l'applicazione di questo strumento complessa; tuttavia, i maggiori ostacoli provengono dall'informazione. In particolare, tali ostacoli si annidano nel concetto di danno e nel nesso di causalità.

- ***Individuazione dei danni compensabili***.

Nella complessità del mondo reale le esternalità raramente rappresentano un fenomeno circoscritto. Gli effetti nocivi procurati, ad esempio, da una sostanza tossica sul sistema nervoso di un individuo possono costituire un danno alla salute, in senso stretto, e poi un danno in termini di perdita di giorni di lavoro, un danno psicologico, un danno morale. Analogamente, l'impatto di un oil spill è scindibile in una pluralità di danni che vanno da quello arrecato all'ecosistema a quelli procurati alla pesca e al turismo. A loro volta, in molti casi, le diverse classi di danni possono essere scomposte in sottoclassi, e così via. Generalmente, non esiste un danno ambientale unico, piuttosto un insieme di "micro-danni" appartenenti a classi diverse che, insieme, costituiscono il danno ambientale complessivo. La valutazione d'impatto ambientale non è altro che un esercizio di ricostruzione capillare dei diversi, singoli effetti sull'ambiente che costituiscono tale danno complessivo. Ora, al di là del problema dell'attribuzione dei pesi ai singoli danni e della loro aggregazione, il ricorso alla responsabilità civile per i casi di danno all'ambiente pone un problema di individuazione delle classi di danno da compensare e, dunque, delle classi di beni che devono essere protette: per quali tipi di danno vi è colpevolezza e/o responsabilità dell'injurer? Nel caso di danno morale, ad esempio, si deve ritenere l'injurer responsabile? E un danno all'ambiente, indipendente da effetti nocivi

sulle persone, deve essere compensato? E se sì, come deve avvenire la compensazione, a quale agente deve essere versato l'indennizzo? Per meglio rispondere a queste domande è bene operare una semplificazione e dividere i diversi tipi di danno in tre grandi classi tradizionali: a) danni alle cose; b) danni alle persone; c) danni all'ambiente. Ora, sul piano economico non si pongono problemi di distinzione tra questi tre tipi di danno. L'obiettivo dell'efficienza richiede che esternalità negative a cose, persone o, più in generale, all'ambiente, debbano in ogni caso essere internalizzate e, pertanto, è necessario che chi procura il danno ne sostenga anche il costo. Tuttavia, esiste un problema di adeguamento della realtà a tale stringente conclusione economica. Ciò è dovuto al fatto che lo strumento della responsabilità civile non nasce con obiettivi di tutela dell'ambiente. Al contrario, tali obiettivi costituiscono un esigenza recente: in Italia, ad esempio, la responsabilità civile, come prevista dal codice civile, è stata concepita per la tutela di diritti individuali, valutabili economicamente e diversi da quelli ambientali[31]. Ed anche negli altri paesi, la responsabilità civile rappresenta uno strumento nato con finalità diverse da quelli di protezione dell'ambiente e, quindi, essa deve essere rivista e rinnovata onde rendere possibile un'applicazione al contesto dell'ambiente. In particolare, ciò che occorre - e ciò che è stato in parte, e in modi diversi, realizzato - è un'estensione dell'area soggetta a responsabilità civile. Mentre la classe dei danni alle cose è tutelata, da sempre, dallo schema generale della responsabilità civile, lo stesso non può dirsi per i danni alle persone e all'ambiente. In Italia, fino a non molto tempo fa, la compensazione per danni alla persona era prevista solo per la parte riguardante gli effetti economici negativi del danno fisico sul patrimonio della vittima. Il danno biologico in sé, ovvero la compromissione della salute psico-fisica dell'agente, non era riconosciuto. Ma è la classe dei danni all'ambiente quella che ha posto, e pone, i maggiori problemi di adeguamento. L'ambiente deve essere ritenuto un bene giuridico autonomo, oppure, al contrario, si deve ritenere i danni all'ambiente come afferenti alle classi tradizionali dei danni alle cose e alla persone? Qui, si è di fronte ad un ostacolo di natura essenzialmente tecnica: il problema non è quello riguardante la necessità, o meno, di estendere la responsabilità civile al contesto dei danni ambientali. Piuttosto, posto che tale estensione è necessaria per la tutela dell'ambiente, essa come deve essere realizzata? Non

[31] Cfr. Rapisarda Sassoon C., 1990, p. 99.

esiste una risposta univoca: i diversi ordinamenti hanno proposto soluzioni differenti. In particolare, si possono distinguere tre modelli di soluzione[32]:

- **l'ambiente non viene riconosciuto come bene giuridico autonomo**. Un esempio di tale impostazione è costituito dalla legge tedesca del 1991. Essa afferma: "qualora da un'immissione nell'ambiente, proveniente da uno degli impianti indicati nell'appendice 1 (della presente legge), derivi la morte, la lesione del corpo o della salute, oppure della proprietà di un soggetto, il proprietario dell'impianto è tenuto a risarcire il danno che ne è derivato" . Nel modello tedesco, dunque, non si fa menzione di un danno diretto arrecato all'ambiente, al contrario i danni scaturiti da immissioni nocive nell'ambiente vengono ricondotti alle classi dei danni alle persone e alle cose (proprietà, patrimonio);

- **l'ambiente viene riconosciuto come bene giuridico autonomo e vengono esplicitate le azioni che sono da ritenersi nocive**. Ad esempio, la legge italiana 349 del 1986 afferma che "è compito del Ministero assicurare... la promozione, la conservazione ed il recupero delle condizioni ambientali conformi agli interessi fondamentali della collettività ed alla qualità della vita, nonché la conservazione e la valorizzazione del patrimonio naturale nazionale e la difesa delle risorse naturali dell'inquinamento" (art. 1). D'altra parte, questa estensione della tutela dello stato ai beni ambientali e le conseguenti potenzialità di internalizzazione della legge vengono contrastate nei fatti da un vincolo posto sul tipo di violazione riconosciuta. L'articolo 18 della 349, infatti, dispone che lo stato può richiedere risarcimenti per danni all'ambiente solo nei casi in cui il danneggiante violi, dolosamente o colposamente, disposizioni di legge o provvedimenti adottati in base a legge. Non solo si opta per un regime di responsabilità per colpa, ma si fa dipendere il risarcimento dalla violazione di disposizioni di legge e di provvedimenti già adottati. Il CERCLA (Comprehensive Environmental Response, Compensation and Liability Act) statunitense del 1980, che costituisce ormai un famoso punto di riferimento per la legislazione

[32] Tale distinzione in tre modelli di soluzione è tratta da Pozzo B., 1994, "Il criterio di imputazione della responsabilità per danno all'ambiente nelle recenti leggi ecologiche", p. 11, in Trimarchi P. (a cura di), 1994. A questo testo si rimanda per approfondimenti sui diversi modelli nazionali e sui problemi di imputazione della responsabilità civile all'ambiente.

ambientale, rappresenta un altro esempio di riconoscimento dell'ambiente come bene giuridico autonomo. In particolare, lo stato viene considerato amministratore delle risorse naturali le quali vengono definite come il complesso di: "land, fish, wildlife, biota, air, water, ground water, drinking water supplies, and other such resources belonging to, managed by, held in trust by, appartaining to, or otherwise controlled by the United States... any State or local government, or any foreign Government"[33]. Le risorse naturali, dunque, vengono poste dal legislatore americano sullo stesso piano di qualsiasi altra risorsa e, dunque, azioni nocive ai loro danni devono essere internalizzate. Inoltre, nella legge americana si esplicita l'insieme dei danni all'ambiente che verranno puniti: "natural resources damages resulting from a release of a hazardous substance";

- **l'ambiente viene riconosciuto come bene giuridico autonomo ma non vengono esplicitate le tipologie di azione nociva**. La "legge base sull'ambiente" del Portogallo (7 aprile 1987) è un esempio di tale impostazione: "ogni cittadino ha diritto ad un ambiente umano ed ecologicamente equilibrato ed ha il dovere di difenderlo", dove per ambiente s'intende "l'aria, la luce, l'acqua, il suolo ed il sottosuolo, la flora e la fauna". Inoltre, l'articolo 41 dispone che "l'autore di danni significativi all'ambiente, derivanti da attività particolarmente pericolosa è tenuto a ripararli indipendentemente da qualsiasi colpa, anche in assenza di una violazione di una norma in vigore". Pertanto, in questo caso si ha, da un parte, un riconoscimento esplicito del regime della responsabilità oggettiva, dell'ambiente come bene giuridico e della necessità di proteggerlo, mentre, dall'altra, non si vincola l'obbligo della compensazione alla violazione di norme già in vigore.

Questi tre casi, che sintetizzano possibili approcci giuridici all'ambiente, documentano l'esistenza di gradi di libertà che il legislatore ha nel trattamento delle esternalità ambientali. Il primo caso può essere considerato la soluzione più restrittiva e, in certa misura, antropocentrica, poiché il danno acquisisce rilevanza solo se nuoce all'uomo, alla sua salute o al suo patrimonio. Il secondo caso, al contrario, rappresenta uno scostamento

[33] In Pozzo B., 1994, p. 13.

dall'antropocentrismo poiché i diversi elementi che costituiscono l'ambiente (l'aria, l'acqua, il suolo, ecc.) vengono considerati in se stessi e si esplicita l'esigenza di una loro protezione. Infine, il terzo caso, non solo ribadisce la prospettiva biocentrica già presente nel secondo ma estende anche la necessità della riparazione a tutti i danni significativi all'ambiente. Naturalmente, questa interpretazione dei tre casi ha il limite di prendere in considerazione unicamente - e parzialmente - l'aspetto normativo formale dei diversi ordinamenti e di trascurare gli aspetti di attuazione della norma. Apparati pubblici inefficienti e oscurità e involuzione nel complesso di leggi che corredano le norme generali possono inficiare il principio astratto e tradurre lo sforzo del legislatore in velleitaria costruzione verbale.

• ***Individuazione del nesso di causalità.***

Nel mondo reale, il legame causale tra danneggiante e danno è spesso oscuro, confuso, poco evidente. Risalire dal danno al danneggiante è, in alcuni casi, un'operazione complessa o, addirittura, impossibile. Ciò per numerose ragioni.

- **Pluralità di cause**. Così come un singolo incidente ambientale spesso causa una molteplicità di danni, un singolo danno può essere l'effetto di più cause: le aziende A, B, e C - che sviluppano il medesimo processo produttivo - possono contribuire, ma in misura diversa, alla generazione del danno Z. Oppure, in condizioni normali la loro attività non genera danni, tuttavia, il verificarsi di un incidente - ad esempio, in B - procura un danno alla salute della popolazione dei centri urbani vicini alle tre aziende. In questi casi, come imputare la responsabilità alle tre parti? Se le tre aziende emettono nell'atmosfera la stessa sostanza, come individuare il loro contributo al danno? E nel caso in cui il danno sia stato causato da una sola delle tre, ma non è possibile dimostrarlo, come procedere[34]? Tale possibilità di pluralità delle cause raggiunge la sua maggiore manifestazione nel caso dei fenomeni ambientali globali. Fenomeni quali le piogge acide, il deterioramento dello scudo d'ozono ed il warming coinvolgono più stati - e dunque danneggiano una moltitudine di persone - e sono causati da un numero elevatissimo di agenti. C'è un effetto ma la causa è

[34] Per approfondimenti sui problemi connessi all'identificazione del nesso di causalità, Villa G., 1994, "Nesso di causalità e responsabilità civile per danni all'ambiente", in Trimarchi P. (a cura di), 1994.

"lontana", ovvero il contributo di ogni singolo agente al danno è minimo. Lo schema della responsabilità civile, in questi casi - per ragioni concernenti anche la sovranità nazionale, le differenze tra ordinamenti giuridici, le relazioni tra stati, i diversi interessi, ecc. - non può essere applicato.

- **Danno sinergico**. L'azione congiunta di A, B e C - che sviluppano processi produttivi diversi - può essere la causa di un danno maggiore di quello risultante dalla somma dei danni che le tre aziende avrebbero potuto provocare isolatamente. In questo caso, bisogna ritenere ciascuna delle tre aziende responsabili di un danno maggiore di quello che ciascuna di essa provocherebbe da sola? Oppure: l'azione congiunta di A, B e C, ossia l'interazione delle loro emissioni, è la causa di un danno che non si sarebbe verificato se A, B e C avessero operato isolatamente. In altri termini, A, B e C non sono inquinanti ma la combinazione dei loro processi lo è. In questi caso, vi è una responsabilità da parte delle tre aziende? E se sì, in che misura?

- **Ritardo temporale**. Il danno provocato da A, B e C - isolatamente o in congiunzione - si manifesta molti anni dopo lo svolgimento dell'attività produttiva, eventualmente quando essa è già cessata. Per quanto tempo A, B e C sono soggette a responsabilità? E se la loro attività è cessata - ed esse, come persone giuridiche non esistono più - a chi deve essere imputata la responsabilità?

- **Danno probabile**. La vittima, che vive nell'area in cui operano A, B e C, contrae una malattia cancerogena. Le emissioni chimiche di A, B e C producono un rischio di danno, ovvero vi è la possibilità che esse siano la causa del cancro nella vittima. Tuttavia, non vi è certezza: come dimostrare che il tumore sia stato causato proprio dalle emissioni di A, B e C e non, ad esempio, da alcolismo, fumo, esposizione ad altri agenti inquinanti in altri luoghi, tipicità genetica della vittima?

- **Causa nascosta**. Un azienda zootecnica collocata a monte inquina, ad intervalli irregolari, un fiume e danneggia un agricoltore collocato a valle che usa l'acqua del fiume per irrigare le sue terre. Se vi fosse un sistema di monitoraggio efficiente la responsabilità dell'azienda emergerebbe, ma poiché il monitoraggio è insufficiente la causa rimane non identificata.

Questi casi riassumono le principali difficoltà che è possibile riscontrare nell'individuazione dei nessi di causalità. Il regime della responsabilità oggettiva, per quanto più semplice rispetto a quello di responsabilità per colpa, ne risulta anch'esso ostacolato. I diversi ordinamenti, comunque, pur non offrendo risposte puntuali a ciascun problema, propongono - in modo esplicito o implicito - alcuni modelli di soluzione. In particolare, si possono individuare due differenti impostazioni:

- **modello statunitense**: responsabilità proporzionale. Nel caso di in danno provocato da più agenti, la responsabilità è imputata alle diverse parti in misura proporzionale alle quote di mercato che ciascuna di essa detiene. Un esempio è quello costituito dalla causa Sindell v. Abbot Laboratories. Negli USA tra il 1941 ed il 1971, alcune case farmaceutiche offrirono sul mercato un prodotto, contenente un estrogeno sintetico (il dietilstilbestrolo), per la prevenzione di interruzioni di gravidanza dovute a insufficienze ormonali. Intorno alla fine degli anni '60, alcuni studiosi dimostrarono che il dietilstilbestrolo causava tumori in donne nate da madri che avevano utilizzato il prodotto in questione. Così, nel 1971, il prodotto venne escluso dal commercio. Numerose vittime chiamarono in giudizio le aziende farmaceutiche, tuttavia molti anni erano passati da quando tali aziende avevano cominciato a distribuire sul mercato il dietilstilbestrolo: come attribuire la responsabilità alle oltre duecento aziende che lo producevano? Quali di esse aveva venduto il prodotto - ad esempio - a Y, da cui era nata X, che si era ammalata di cancro? La Corte Suprema della California, dovendo decidere di questo problema, stabilì che ciascuna azienda farmaceutica che anni prima aveva offerto il prodotto avrebbe potuto essere responsabile. In particolare, la Corte decise che il grado di responsabilità delle singole case farmaceutiche - che riflette la probabilità di causare la malattia - era proporzionale alla

quantità di prodotto offerto sul mercato, ovvero alla quota di mercato della singola azienda. E' questo il principio della *market share liability* che attribuisce la responsabilità al singolo produttore in relazione alla quota di mercato da esso detenuta. Per un verso, tale principio costituisce uno strumento di internalizzazione dei costi sociali perché c'è da ritenere che, in media, ogni produttore sarà ritenuto responsabile della quota di danno totale da esso causata. Per un altro verso, però, tale principio risulta indebolito dall'assunzione - sulla quale implicitamente poggia - di omogeneità dei prodotti, indebolimento che ne ostacola una trasposizione rigida all'ambiente: perché, ad esempio, due aziende che detengono ciascuna il 50% della quota di mercato di una certa produzione - ma che inquinano in misura diversa poiché impiegano tecnologie differenti - devono essere ritenute responsabili in pari misura dei danni causati dalle loro emissioni inquinanti? Problemi come questo rendono il principio della *market share liability* poco robusto per applicazioni in campo ambientale in tutti quei casi in cui l'insieme dei potenziali danneggianti è caratterizzato da una forte eterogeneità;

- **modello tedesco**: responsabilità solidale. Se, in caso di danno causato da più agenti, l'individuazione dei responsabili non è possibile, ogni potenziale danneggiante è ritenuto responsabile del danno totale, e naturalmente può, in seguito, rivalersi sugli altri responsabili. Certamente ciò incentiva la sicurezza perché il potenziale danneggiante è comunque, in prima battuta, responsabile e, dunque, sarà indotto alla diligenza. D'altra parte, questo principio, adottando la presunzione di responsabilità nei casi incerti, può rivelarsi eccessivamente restrittivo verso coloro che svolgono attività rischiose. Nel diritto tedesco, sono state così adottate alcune disposizioni che mitigano la presunzione di responsabilità del potenziale injurer e rendono il principio della responsabilità solidale più elastico, avvicinandolo a quello della responsabilità proporzionale[35].

[35] E' il caso del concorso di colpa da parte della vittima, oppure di danno arrecato da incapace. Per approfondimenti, cfr. Villa G., 1994, op. cit., dove viene esaminato anche un terzo modello - francese, denominato "perdita della chance" - di imputazione della responsabilità.

In sintesi, questi due modelli, tentano di risolvere un problema tecnico: posto che l'obiettivo è l'attribuzione di un costo sociale a colui che lo causa, e posto che alle volte vi è incertezza e non è possibile conoscere l'effettivo injurer, come procedere? Il primo modello costituisce certo una risposta più puntuale perché tenta di individuare - pur nell'approssimazione derivante dal riferirsi a quote di mercato - il contributo di ogni singolo potenziale danneggiante. Il secondo modello, invece, non compie questo tentativo. La sua applicazione, pertanto, è più semplice poiché non si richiedono informazioni statistiche per risalire al grado di responsabilità di ogni singola fonte inquinante. E in certa misura, dunque, vi è un guadagno in semplicità conquistato con una flessione della precisione.

- ***Valutazione del danno.***

Nello schema astratto ad ogni tipo di danno corrisponde una misura monetaria definita che consente la compensazione, il calcolo economico da parte degli agenti, l'emissione di un giudizio di negligenza o di diligenza da parte della Corte. Nella realtà, soprattutto per certi tipi di danno, si è assai lontani dalla definizione univoca di una misurazione monetaria. Ciò pone numerosi problemi, da quelli tecnici di individuazione del danno monetario arrecato all'ambiente e della somma da versare alla vittima (quanto vale una malattia polmonare causata dalle emissioni nocive di un cementificio o di un impianto chimico? E la morte di 100.000 pesci di acqua dolce? E la distruzione di 300 ettari di bosco?) a quelli concernenti il calcolo economico delle autorità e degli agenti. L'informazione sull'entità del danno rappresenta un elemento essenziale del calcolo di efficienza: se il danno non può essere misurato univocamente diviene improbabile sia la definizione di uno standard ottimale da parte delle autorità sia l'ipotesi di un comportamento improntato all'efficienza da parte degli agenti. La letteratura sulla valutazione del danno e, più in generale, delle esternalità è ingente[36]. Al di là delle caratteristiche delle numerose tecniche proposte, qui è rilevante porre l'attenzione su tre aspetti cruciali: i) l'entità del danno, che nello schema astratto viene assunta come data e come fondante il calcolo economico degli agenti, costituisce in realtà, e in molti casi, un elemento

[36] Negli ultimi anni, una notevole mole di articoli e libri - di divulgazione e di approfondimento teorico e tecnico dell'argomento - è stata proposta all'attenzione degli studiosi di ambiente, degli economisti applicati e degli operatori pubblici. Un testo di sintesi dell'argomento - uno per tutti - è Freeman III, Myrick A., 1993.

piuttosto sfuggente; ii) l'intera efficacia dei regimi di responsabilità e degli strumenti di politica ambientale è subordinata all'adozione di soluzioni "accettabili" del problema della valutazione di ciò che non ha prezzo. Negli Stati Uniti, ad esempio, dove la tecnica della valutazione contingente (rilevazione della disponibilità a pagare degli agenti attraverso questionario) è largamente impiegata per la stima monetaria delle compensazioni associate a liability, le autorità hanno affidato ad un panel di studiosi, tra i quali alcuni premi Nobel, la redazione di un rapporto che indicasse delle linee guida da seguire nell'uso della tecnica[37]. In altri termini, un impiego della liability orientato a "scientificità" - anche attraverso valutazioni del danno che sfruttano i progressi della ricerca economica applicata - è indissolubilmente legato alla individuazione di robusti e approvati criteri di valutazione delle compensazioni e dei danni; iii) il ricorso a procedure di valutazione delle esternalità collaudate e "scientifiche" comporta un costo. Quanto più cresce l'attenzione posta nella stima delle compensazioni tanto più, verosimilmente, cresce il suo costo. Tale costo, analogamente a quelli di transazione nella contrattazione e quelli delle spese giudiziali, può avere un'incidenza rilevante e, dunque, deve essere tenuto in considerazione nel calcolo economico.

3.6. Tendenze recenti

Negli ultimi anni, i regimi di responsabilità sono stati al centro di un vivace dibattito. Stimolata da una proposta della Commissione CEE[38] - che senz'altro riflette l'interesse degli stati membri intorno al tema - è tuttora in corso una discussione sull'opportunità di utilizzare lo strumento della responsabilità oggettiva per i casi di danno ambientale. Già con la Direttiva 85/374 del 1985 era stata introdotta, in ambito comunitario, la responsabilità oggettiva per i prodotti difettosi, ovvero la responsabilità del produttore di articoli che causano danno ai consumatori, indipendentemente dalla colpa. Tale Direttiva, tuttavia, circoscriveva i danni rimborsabili alla sfera privata e non considerava i casi di danno all'ambiente. Il "Libro Verde" della Commissione CEE costituisce un primo passo verso un impiego armonizzato della

[37] Arrow, Kenneth et al., 1993.

[38] "Il libro verde sul risarcimento dei danni all'ambiente", Bruxelles 14 maggio 1993. Tutte le citazioni fatte nel paragrafo 3.5. sono tratte da tale documento.

responsabilità civile per i danni ambientali. Esso interpreta la responsabilità civile come un mezzo di prevenzione e di applicazione del principio "chi inquina paga" e, più in generale, costituisce un'esplorazione delle caratteristiche dei regimi di responsabilità e dei principali aspetti critici: le diverse tipologie di danno, le difficoltà che esistono nell'imputazione della responsabilità, il nesso di causalità, la legittimazione ad agire, ecc. In sintesi, il documento riassume le condizioni di applicabilità della responsabilità civile per i danni all'ambiente nello schema mostrato alla pagina seguente[39].

Lo studio della Commissione, nato con l'intento esplicito di suscitare un ampio dibattito, costituisce un punto di riferimento importante perché suggerisce, nel quadro piuttosto variegato delle diverse legislazioni nazionali, l'impiego della responsabilità oggettiva per i danni ambientali. Esso, pur riconoscendo l'utilità della responsabilità per colpa, ne ridimensiona il ruolo nel contesto dei danni ambientali, perché:

Condizioni di applicabilità, e non, della responsabilità civile

[39] Adattato dal "Libro Verde" della CEE.

Applicabilità		**Non applicabilità**
• Danno misurabile ed immediato		• Danno latente o indeterminabile
• Atto o incidente circoscritto		• Atti o incidenti cumulativi e non circoscritti
• Parti responsabili identificabili		• Parti responsabili non identificabili
• Responsabilità (per colpa o ...		• .... assenza di base per responsabilità oggettiva)
• Determinabilità del nesso causale		• Nesso causale non determinabile
• Parte legittimata all'azione che abbia interesse ad agire		• La parte legittimata all'azione non ha interesse ad agire
↓		↓
Possibilità dell'azione di responsabilità civile		***Non utilità della responsabilità civile: necessità di uno strumento di risarcimento collettivo***

- "nel caso di danno ambientale non è certo facile fornire la prova della colpa. In taluni casi è addirittura impossibile";

- l'onere della prova spetta alla vittima che deve dimostrare la negligenza dell'injurer. Ma questo non sempre è possibile, anche perché la legislazione ambientale presenta molte lacune sicché una valutazione di quale comportamento costituisca diligenza e quale no, può non risultare agevole.

La responsabilità oggettiva, invece, "appare particolarmente adatta alle caratteristiche di riparazione del danno ambientale", perché:

- rappresenta, rispetto alla responsabilità per colpa, una semplificazione in quanto non è necessario dimostrare la colpa.

- "contribuirà a favorire una migliore gestione dei rischi e rappresenterà una sicurezza giuridica per le aziende coinvolte". Essa "può anche favorire l'attuazione del principio chi inquina paga per determinati tipi di attività economica".

La Commissione, comunque, è piuttosto cauta nella sua valutazione positiva della responsabilità oggettiva e segnala l'esigenza di superare alcuni ostacoli affinché lo strumento in questione possa essere pienamente operativo. In particolare, "l'intero dispositivo deve essere soggetto a definizioni molto precise". Occorre definire giuridicamente cosa significhi danno ambientale, ovvero "il grado di degrado a partire dal quale si considera avvenuto un danno"; poi, occorre individuare le attività alle quali si applica il regime della responsabilità oggettiva, cercando di restringere sufficientemente il campo, poiché "un sistema troppo esteso, comprendente un numero troppo elevato di attività, rischia di avere conseguenze negative per gli operatori economici e di creare un clima di incertezza giuridica che lo rende inapplicabile"; infine, è necessario definire la persona o la parte cui incombe la responsabilità, essendo possibile imputare la responsabilità a più di un agente, ovvero "alla persona in possesso delle conoscenze tecniche, a quella in possesso delle risorse o del controllo operativo dell'attività". In generale, a tali problemi tecnici di attuazione della norma, si aggiungono ostacoli di carattere politico. Le tradizioni degli stati membri in tema di responsabilità civile sono piuttosto variegate e, a volte, all'interno di uno stesso ordinamento possono coesistere approcci differenti[40]. Ciò rende l'armonizzazione degli ordinamenti, obiettivo esplicitato anche dal "Quinto programma d'azione ambientale" della Comunità, un obiettivo piuttosto ambizioso. Inoltre, l'atteggiamento delle diverse parti sociali non è omogeneo. Da una parte, i gruppi industriali hanno reagito negativamente alla proposta di un'adozione diffusa della responsabilità oggettiva a ragione degli oneri che essa farebbe gravare sul mondo dell'industria, dall'altra parte, i gruppi ambientalisti hanno accolto favorevolmente la proposta CEE e sostenuto l'ipotesi di retroattività della

[40] Per una panoramica sui diversi regimi di responsalità in Europa, Brealey, Mark 1993.

responsabilità e di rovesciamento dell'onere della prova, con l'accusato - il presunto injurer - che deve dimostrare la propria innocenza.

In ambito europeo, comunque, il problema del risarcimento per i danni all'ambiente è stato affrontato anche dalla Convenzione di Lugano. Il testo definitivo (marzo 1993) si concentra sui danni ambientali derivanti da attività pericolose (attività concernenti sostanze pericolose, microrganismi, trattamento e smaltimento rifiuti) e adotta lo strumento della responsabilità oggettiva. A differenza del Libro Verde, che costituisce un documento quadro sul tema del risarcimento del danno ambientale in generale, non solo da attività pericolosa, la Convenzione ha valore vincolante per le parti firmatarie[41].

Fuori dell'Europa, particolarmente interessanti sono la legislazione giapponese e quella statunitense. La prima adotta il regime di responsabilità oggettiva per i casi danni fisici provocati da inquinamento atmosferico ed idrico; per i danni non fisici, invece, occorre dimostrare la colpevolezza di chi ha inquinato. Inoltre, il sistema giapponese si fonda sulla "teoria del limite sopportabile" e "sulla teoria della probabilità". Per la prima teoria, il danneggiato può adire il tribunale solo quando il danno supera il limite sopportabile. Per la seconda, ciò che la vittima deve dimostrare è solo la probabilità che esista un legame di causalità tra azione e danno. Infine, un fondo di indennizzo, finanziato dai contributi sulle emissioni inquinanti e da una quota dell'imposta sui veicoli a motore, assicura alla vittima, in casi specificati dalla legislazione, una compensazione anche quando il responsabile o il legame di causalità sono incerti.

La legislazione statunitense, invece - come abbiamo già visto -, si fonda sul CERCLA del 1980 ed adotta la responsabilità oggettiva. Il Superfund, qualora i danneggianti non fossero identificabili o non adottassero le azioni necessarie, dovrebbe consentire azioni rapide di pulizia e di risanamento mentre il governo, in seguito, può recuperare le spese di ripristino presso le

[41] I primi paesi a firmare sono stati: Cipro, Finlandia, Grecia, Islanda, Italia, Liechtenstein, Lussemburgo, Olanda.

parti potenzialmente responsabili[42]. Queste ultime sono responsabili non solo per i costi di ripristino di un sito ma anche per i danni alle risorse naturali. Inoltre, i cittadini possono esercitare azione legale nei casi un cui l'autorità pubblica è inadempiente e non svolge il suo ruolo di protezione dell'ambiente. Nell'ottobre 1986 il CERCLA è stato integrato dal SARA (Superfund Amendment and Reauthorisation Act) che, oltre ad aumentare il fondo fiduciario da 1.6 ad 8.5 miliardi di dollari, introduce una possibilità di difesa dell'accusato oltre quelle previste dal CERCLA (an act of god, an act of war, an act or omission of a third party..., any combination of the foregoing): la "innocent landowner defence", infatti, consente all'accusato di essere scagionato qualora egli non avesse rilevato, dopo un'adeguata valutazione precedente la transazione per l'acquisizione del sito, che esso era contaminato. In realtà, però, il CERCLA si è dimostrato macchinoso e costoso dando luogo ad un elevato numero di litigi, spesso complessi e coinvolgenti un ampio numero di agenti: "il numero dei litigi e delle procedure in corso rappresenta circa il 30-60% delle risorse spese dall'agenzia per la protezione dell'ambiente e dagli operatori e assicuratori interessati".

In generale, dunque - come l'esperienza statunitense conferma -, l'applicazione dello strumento della responsabilità civile per i danni ambientali, per quanto promettente in astratto, appare nel concreto ostacolata da frizioni che lo sforzo razionalizzatore delle autorità deve ancora eliminare. Inoltre, l'armonizzazione delle legislazioni dei diversi stati - in misura tale da eliminare, o ridurre, effetti di distorsione della concorrenza - si scontra con l'eterogeneità delle tradizioni giuridiche e con il contrapposto atteggiamento che i diversi gruppi d'interesse hanno nei confronti dell'ambiente. Analogamente ad altri strumenti della politica ambientale, parte della forza della responsabilità civile si dissìpa negli attriti del reale.

[42] Parti potenzialmente responsabili possono essere: il proprietario attuale del sito, il proprietario del sito nel momento in cui avvenne la contaminazione, l'operatore industriale che ha generato lo scarto, il trasportatore dello scarto, il negoziante dello scarto.

4. L'ASSICURAZIONE E IL PROCESSO

4.1. L'assicurazione

L'influenza che i regimi di responsabilità hanno sui comportamenti dipende anche dagli atteggiamenti che gli agenti hanno nei confronti del possibile incidente. Poiché si tratta di un evento incerto - che potrebbe verificarsi o meno nel futuro - gli agenti non fronteggiano un fatto reale e, pertanto, risarcimenti e danni causati o subiti sono - ex ante - soltanto ipotetici: gli agenti non si confrontano con una somma monetaria certa ma, come abbiamo visto, con un valore atteso. Ora, tale situazione fa sì che l'atteggiamento che gli agenti hanno nei confronti dell'incidente - in particolare, la loro attitudine rispetto al rischio - divenga un elemento rilevante dell'indagine. Un danno certo di 1.000 $ vale, nel conto economico di qualsiasi agente, 1.000 $. Lo stesso non può dirsi di un danno atteso: 1.000 $ di valore atteso possono, come somma di denaro, essere percepiti in modo diverso dai diversi agenti. Così, per un agente neutrale rispetto al rischio 1.000 $ di valore atteso equivalgono a 1.000 $ certi, mentre per agenti avversi o propensi al rischio questa equivalenza non vige più. In altri termini, solo se l'agente è neutrale rispetto al rischio è lecito assumere che egli baserà i propri calcoli economici sul valore atteso, e preverrà ogni qual volta il costo della prevenzione è inferiore al costo atteso dell'incidente, e viceversa. Qualora, invece, l'agente è propenso al rischio questo risultato potrà non essere valido: l'agente, ora, poiché è incline al rischio in sé, potrà non prevenire l'incidente preferendo rischiare di avere nel futuro una forte, ma incerta, perdita. Al contrario, un agente avverso al rischio, preferirà sempre l'alternativa certa, ovvero preferirà la prevenzione anche qualora il suo costo sia solo pari - non occorre che sia inferiore - al danno atteso[43].

[43] Questi risultati, in verità, sono messi in discussione dall'approccio di Kahneman e Tversky fondato sulla funzione asimmetrica del valore. Questi studiosi documentano attraverso esperimenti come il modello di scelta tradizionale fondato sull'utilità attesa possa essere, in certe situazioni, sbagliato. Così, secondo il modello di Kahneman e Tversky, un agente avverso al rischio sceglierà, tra una perdita certa e una perdita attesa di uguale ammontare, la seconda alternativa, ovvero l'opposto di quanto previsto dalla teoria tradizionale. Qui, dati gli obiettivi di queste note, tali scostamenti dal modello tradizionale non verranno presi in considerazione.

Inoltre, occorre tener conto del fatto che i due regimi della responsabilità oggettiva e della responsabilità per colpa allocano il rischio tra i due agenti - injurer e vittima - in modo asimmetrico.

- In un regime di responsabilità oggettiva, l'injurer è tenuto al risarcimento del danno ogni qual volta esso si verifica, e ciò significa che la vittima - assumendo perfetta solvenza dell'injurer - è sempre compensata e, dunque, di fatto, assicurata. In questo caso, il rischio d'incidente grava interamente sull'injurer perché è lui che ne sostiene totalmente il costo.

- Nel caso della responsabilità per colpa, invece, se l'injurer è stato diligente ed ha adottato tutte le misure di prevenzione, egli non è ritenuto responsabile e, pertanto, non deve indennizzare la vittima. La vittima, in questo caso, non è compensata e, quindi, non è assicurata: il rischio grava interamente su di lei.

Data questa distribuzione asimmetrica del rischio tra gli agenti, il regime della responsabilità oggettiva rappresenterà lo schema ideale di allocazione del rischio in tutte quelle situazioni nelle quali l'injurer è neutrale rispetto al rischio e la vittima è avversa al rischio. Viceversa, per i casi nei quali l'injurer è avverso al rischio e la vittima è neutrale, il regime di responsabilità per colpa costituirà la soluzione ottimale. Naturalmente, quando entrambi gli agenti sono avversi al rischio, occorrerà calibrare una soluzione intermedia, ad esempio un regime di responsabilità oggettiva "mitigato" nel quale l'injurer compensi la vittima solo per una parte del danno. D'altro canto, occorrerà tenere presente che tale indebolimento sostanziale della responsabilità dell'injurer ha un costo in termini di prevenzione perché, quanto più diminuisce l'indennizzo versato dal danneggiante al danneggiato, tanto più il primo sarà disincentivato a prevenire[44].

Ora, nella realtà - ed entro certi limiti, come vedremo in seguito -, gli agenti hanno la possibilità di assicurarsi contro eventuali incidenti. Ciò significa che essi possono proteggersi dal rischio stipulando una polizza di assicurazione che li copre dal rischio, rimuovendolo del tutto. A

[44] Per approfondimenti istituzionali sugli aspetti di attitudine al rischio, comportamento degli agenti, regimi di responsabilità e assicurazione, Polinsky, M. A. 1989, cap. 7, 9, 10.

fronte del pagamento certo di un premio di assicurazione monetario, gli agenti possono coprirsi contro il rischio perché, nel caso in cui l'evento incerto - incidente - si verificasse, l'assicurazione provvederebbe alla compensazione delle vittime. In un mondo ideale nel quale non esistessero costi amministrativi e le compagnie di assicurazione fissassero i premi al livello sufficiente a coprire i rimborsi per incidenti, tali premi rifletterebbero il danno atteso. Un agente avverso al rischio, sul quale grava il rischio di incidente, date queste ipotesi, troverà sempre conveniente stipulare un contratto di assicurazione che rimuove il rischio. Idealmente, avremo che:

- in una situazione nella quale sia vigente il regime della responsabilità oggettiva, il potenziale danneggiato non avrà convenienza ad assicurarsi poiché l'indennizzo da parte dell'injurer costituirà, come abbiamo visto, una forma di assicurazione. Al contrario, il potenziale danneggiante, se avverso al rischio, troverà certamente conveniente stipulare una polizza di assicurazione poiché è su di lui che grava l'onere del rischio;

- in un contesto di responsabilità per colpa, invece, il potenziale danneggiante potrebbe stipulare un'assicurazione contro danni a terzi solo qualora egli scegliesse di comportarsi negligentemente e, dunque, di procurare danno ed essere ritenuto responsabile. Tuttavia, questa è solo una possibilità perché il potenziale injurer, poiché assunto razionale, effettuerà un confronto tra beneficio a lui derivante dall'essere negligente e premio di assicurazione da pagare, premio che riflette - come abbiamo visto - il danno atteso. Ora, se le autorità hanno stabilito lo standard di diligenza in modo efficiente, il potenziale injurer avrà convenienza a rispettarlo e, dunque, a non assicurarsi. La vittima potenziale, d'altra parte, poiché sostiene l'onere del rischio - e, in caso di diligenza dell'injurer, non sarà rimborsato - troverà conveniente, se avverso al rischio, stipulare una polizza di assicurazione contro danni da lui subiti.

Le combinazioni, naturalmente, sono più numerose perché l'avversione al rischio è solo una delle tre possibilità di attitudine al rischio. In un mondo di agenti propensi al rischio, ad

esempio, il quadro cambia e l'assicurazione può diventare superflua; d'altro canto, se tutti gli agenti fossero neutrali rispetto al rischio, vi sarebbe indifferenza tra l'assicurarsi e il non assicurarsi. La realtà, inoltre, è resa più complessa e sfaccettata perché, rispetto allo schema ideale di assicurazione fin qui delineato, esistono delle imperfezioni. Le due imperfezioni tradizionali di un contratto di assicurazione sono il rischio morale (moral hazard) e la selezione avversa (adverse selection):

- si ha rischio morale quando l'agente che si è assicurato ha meno incentivo - poiché il suo rischio è coperto dalla compagnia di assicurazione - ad evitare l'incidente. In altre parole, il fatto che non sia lui a dover compensare i terzi per il danno causato e a rispondere delle sue conseguenze, lo rende meno attento indebolendo o eliminando la sua prevenzione;

- si ha selezione avversa quando agenti ad elevato rischio, che hanno maggiore probabilità di causare incidenti, hanno anche maggiore probabilità di stipulare un contratto di assicurazione. Ossia è plausibile che, ad assicurarsi, siano in larga misura proprio quegli agenti che, poco diligenti o non in condizione di attivare misure di prevenzione adeguate, sono da considerarsi ad elevato rischio.

La conseguenza del moral hazard è che l'assicurazione, poiché abbassa l'incentivo alla prevenzione, può accrescere la probabilità e l'entità dei danni, a svantaggio dell'efficienza. L'adverse selection, invece, può costituire una discriminazione nei confronti degli agenti a rischio medio-basso perché i premi fissati dalle compagnie di assicurazione riflettono la probabilità media di incidente e, dunque, sono tanto più elevati quanto maggiore è il numero di agenti ad elevato rischio che stipulano una polizza di assicurazione. Nel caso del moral hazard è l'assicuratore a correre il rischio di assicurare un agente disincentivato alla prevenzione. Nel caso dell'adverse selection è l'assicurato a rischio medio-basso a dover pagare un premio maggiore di quello che riflette la sua, personale, probabilità di incidente. Queste due imperfezioni, che sono l'effetto di un deficit informativo, potrebbero essere superate qualora le compagnie di assicurazione disponessero di informazione perfetta sugli assicurati. Nel caso del

moral hazard, le compagnie potrebbero aumentare il premio proporzionalmente al decremento di prevenzione dell'assicurato. Nel caso dell'adverse selection, esse potrebbero fissare premi ad hoc, proporzionali alla dose di rischio di ogni assicurato. Occorre osservare, a questo proposito, che nella realtà le compagnie di assicurazione cercano di acquisire informazioni sui propri assicurati e di differenziare i premi, ad esempio esse monitorano le tecnologie di prevenzione degli impianti coperti da polizza, oppure, nel caso del settore automobilistico, incrementano i premi ogni volta che l'assicurato è coinvolto in un incidente (pratica "bonus-malus"). Tuttavia, è nella natura delle cose che la realtà si discosti dal modello ideale e che le compagnie di assicurazione, quindi, non riescano a fissare premi che riflettano perfettamente ed in ogni momento il grado di rischio cui è esposto ogni assicurato. Inoltre, occorre considerare che le caratteristiche e le tipicità del bene ambiente costituiscono un elemento di scostamento ulteriore della realtà dal tipo ideale fin qui esaminato.

4.2. L'assicurazione e l'ambiente.

La copertura del rischio degli agenti è solo una delle funzioni dell'assicurazione. Sul piano analitico, è vero, assicurarsi corrisponde a trasferire il rischio sull'assicuratore. Tuttavia, se spostiamo l'attenzione dal singolo agente al contesto macro e consideriamo più da vicino le tendenze recenti, scopriamo che l'assicurazione, in particolare nel settore ambientale, svolge almeno due altre importanti funzioni. Essa:

- assicura il danneggiato contro eventuali insolvenze dell'injurer in quei casi nei quali il responsabile non è noto, oppure l'injurer, pur se responsabile, non riesce a far fronte al pagamento per debolezza finanziaria. Inoltre, qualora ad essere danneggiato sia semplicemente l'ambiente e vi sia obbligo di ripristino dello stato precedente l'incidente, l'assicurazione - garantendo la copertura - consente un'effettiva protezione dell'ambiente degradato;

- incentiva la prevenzione perché gli assicuratori, adeguando i premi alla dose di rischio dei potenziali danneggianti, li stimolano ad essere più diligenti. Le compagnie di assicurazione, infatti, cercano di proporzionare i premi al rischio, abbassare il massimale di copertura, rifiutare di assicurare clienti particolarmente rischiosi oppure escludere dalla copertura assicurativa specifici rischi[45].

In altri termini, nel settore dell'ambiente, l'assicurazione rappresenta uno strumento che può contribuire a generare efficienza (maggiore prevenzione), giustizia sociale (compensazione delle vittime) e protezione dell'ambiente (garanzia della bonifica e del ripristino dei siti compromessi). Queste funzioni, tuttavia, devono essere considerate potenzialità perché la loro realizzazione non sempre è possibile, essendo essa subordinata al superamento di alcuni ostacoli. In particolare, lo strumento dell'assicurazione risulta indebolito da:

- **incertezza nel nesso di causalità**. In alcuni casi di inquinamento può essere difficile attribuire la responsabilità ad un agente o perché il danno è l'effetto dell'azione di una miriade di agenti (es. fenomeni ambientali globali, come piogge acide, warming, deterioramento dello scudo d'ozono), o perché esso si manifesta in modo occulto o con un notevole lag temporale sicché risulta impossibile risalire alla causa. In casi analoghi, la funzione di indennizzo della vittima è inficiata dalla resistenza che le compagnie di assicurazione fanno ad assicurare i potenziali danneggiati o, se assicurati, a compensarli: l'incertezza nel legame causa-effetto indebolisce lo strumento dell'assicurazione essendo essa idonea in casi di rischi e fenomeni definiti e circoscritti.

- **incertezza nel fenomeno di inquinamento**. Vi sono danni che non vengono considerati tali perché la comunità scientifica non li ha ancora scoperti, danni intorno ai quali non vi è ancora sufficiente consenso scientifico (es. livello di tossicità di determinate sostanze), danni dei quali non sono noti i meccanismi di diffusione. In questi casi, l'insufficiente conoscenza del fenomeno può costituire un limite insuperabile e rendere impossibile

[45] Per approfondimenti sulle funzioni dell'assicurazione ambientale, Cellerino, 1993, cap. 4.

l'assicurazione: la compagnia che assicura può rifiutare di coprire un rischio derivante da un fenomeno incerto, oppure il potenziale danneggiante o la potenziale vittima possono, a causa dell'insufficiente conoscenza del fenomeno, non avvertire l'esigenza di assicurarsi.

- **incertezza nella valutazione del danno**. Si tratta di un ostacolo tipico: il danno all'ambiente o alla persona spesso non può essere valutato attraverso il consueto metro dei prezzi di mercato. In casi semplici nei quali, ad esempio, un lavoratore o un contadino, a causa di un incidente ambientale, perdono giorni di lavoro o il raccolto, è possibile pervenire senza grandi difficoltà ad una stima del danno attendibile e sufficientemente oggettiva perché fondata su valori di mercato (rispettivamente, salario perso e valore del raccolto ai prezzi di mercato). Invece, nel caso di danno, ad esempio, all'ecosistema di un bacino idrico, il processo di valutazione è molto più complesso e può dare luogo a stime non univoche. Tale difficoltà, che rafforza quella della stima del rischio, produce ulteriore incertezza nella definizione del valore atteso del danno. Poiché tale valore è rilevante per la definizione del premio, lo stesso strumento dell'assicurazione ne risulta indebolito.

Questi ostacoli hanno creato, nel settore dell'ambiente, una situazione contraddittoria. Per un verso, lo strumento dell'assicurazione potrebbe svolgere un ruolo rilevante: la maggiore sensibilità ed attenzione verso la protezione dell'ecosistema e degli agenti contro gli effetti esterni fa sì che vengano riconosciute come pericolose, e potenzialmente lesive del benessere, numerose attività dell'uomo e che, quindi, si creino spazi notevoli per uno strumento, quale l'assicurazione, la cui funzione primaria è proprio la copertura del rischio. Per un altro verso, esistono difficoltà "tecniche" nello sfruttamento delle nuove opportunità. La casistica dei danni è estremamente variegata (inquinamenti marini da oil spill, prodotti pericolosi, emissioni dell'industria chimica, nubi radioattive, precipitazioni acide, malattie contratte sul luogo di lavoro, rifiuti pericolosi, erbicidi a rischio, farmaci e sostanze dagli effetti sconosciuti, ecc.) e i danni spesso si scostano dal tipo ideale di danno assicurabile, ovvero un danno derivante da un evento circoscritto nel tempo e nello spazio, effetto certo di una causa ben precisa ed accertabile. Inoltre, a tale complessità si aggiunge la difficoltà intrinseca agli effetti esterni, ossia

l'impossibilità di potere accedere, per essi, ad una valutazione univoca, ad una misura che rappresenti l'esito monetario di uno scambio avvenuto sul mercato. La conseguenza di questo stato di cose è, in primo luogo, che il rischio non viene spesso allocato in modo efficiente. Se gli agenti, a causa degli ostacoli menzionati, non si assicurano o si assicurano in misura insufficiente, essi continueranno a sostenere, in tutto o in parte, l'onere del rischio. D'altra parte, le compagnie di assicurazione possono, a causa di una sottovalutazione del danno atteso, esporsi eccessivamente, ovvero accettare di assicurare rischi ingenti senza un'adeguata copertura, rappresentata dall'incasso di premi elevati. Effetti derivati di questa inefficiente allocazione sono rappresentati dalla compromissione delle altre funzioni che l'assicurazione svolge: l'injurer, poiché non assicurato, può non essere in grado di compensare le vittime; la compagnia di assicurazione, poiché troppo esposta, può fallire; la prevenzione, poiché i contratti di assicurazione sono inefficienti, può indebolirsi e gli incidenti aumentare; le vittime e l'ecosistema possono subire maggiori danni senza che vi sia compensazione o bonifica. Sul piano dell'astrazione economica, il superamento di questi problemi richiederebbe, in primo luogo, un forte investimento in informazione: analogamente agli ostacoli del moral hazard e dell'adverse selection, una crescita della conoscenza, sia da parte degli assicurati che degli assicuratori, estenderebbe il range dei contratti possibili e accrescerebbe la loro trasparenza. Inoltre, i contratti verrebbero orientati verso una maggiore efficienza, evitando esposizioni al rischio imprudenti da parte degli assicuratori. In secondo luogo, a ragione delle specificità dei diversi e numerosi tipi di danno, occorrerebbe differenziazione, flessibilità e trattamenti ad hoc, ovvero miglioramenti organizzativi del settore assicurativo, sostenuti dalla maggiore informazione disponibile.

Ora, le tendenze recenti confermano la rilevanza dei problemi fin qui esposti, che affliggono sia il lato dell'offerta sia il lato della domanda. Per quanto concerne la prima, il caso degli Stati Uniti è emblematico. Come abbiamo visto, il CERCLA impone particolari doveri alle parti potenzialmente responsabili. Esse sostengono i costi delle remedial actions, e inoltre, conformemente al regime di responsabilità, sono tenute al risarcimento dei danni a terzi. I costi, nei casi di danni ambientali particolarmente gravi, possono essere molto elevati. Parecchi

assicuratori statunitensi, essendo elevata la probabilità di compensazioni di dimensione rilevante - ipotesi documentata dalle forti perdite registrate dalle compagnie e da una crisi che ha afflitto il mercato assicurativo dell'America del nord -, hanno rinunciato alla copertura di alcuni rischi ambientali. D'altra parte, sul lato della domanda le imprese spesso sono riluttanti ad assicurarsi. La grande impresa, maggiormente consapevole dei rischi connessi alla propria attività e dotata di maggiori disponibilità finanziarie, si assicura. La piccola e media impresa, invece, meno consapevole e meno ricca, è anche meno sensibile ad assicurarsi. E ciò genera ulteriori problemi, perché vi sono danni ambientali potenziali di grande rilevanza che possono essere generati anche da una piccola impresa; inoltre, la piccola impresa, per la stessa ragione per la quale non si è assicurata - la limitata disponibilità economica - può non riuscire a far fronte al pagamento del danno.

Il mercato assicurativo, dunque, abbandonato a se stesso, può non rispecchiare perfettamente lo schema astratto secondo il quale - come abbiamo visto - date certe condizioni iniziali, gli agenti avrebbero convenienza a coprire il rischio assicurandosi. E' possibile che gli agenti, seppur avversi al rischio, non avvertano l'esigenza di coprirsi oppure, qualora l'avvertano, che non riescano a farlo per ostacoli connessi per lo più all'informazione. In certa misura, quindi, il mercato assicurativo, che costituisce un sotto insieme del mercato globale - ovvero dell'insieme di tutte le interrelazioni dell'economia -, ne eredita i fallimenti e, a sua volta, fallisce nel suo ruolo di generatore dell'efficienza. Così come il mercato globale non riesce a gestire ottimamente le esternalità - a causa degli ostacoli che si oppongono a transazioni compensative tra agenti -, il mercato delle assicurazioni non riesce a gestire ottimamente la loro copertura. In alcuni casi, le autorità pubbliche da una parte, e gli operatori privati dall'altra, hanno introdotto innovazioni che coinvolgono il settore assicurativo e che, in parte ideate come rimedio ad alcuni dei problemi sin qui esaminati, potrebbero contribuire alla loro soluzione. Due sono le linee di tendenza maggiormente rilevanti: l'assicurazione obbligatoria e i fondi di indennizzo.

- L'assicurazione obbligatoria per impianti ad elevato rischio ambientale è stata introdotta in Germania. Inoltre, la stessa Comunità Europea, pur con sfumature, mitigazioni e

consapevolezza della complessità del problema, sembra aderire a questa linea. La proposta di direttiva sulla responsabilità civile per danni all'ambiente causati da rifiuti, presentata dalla Commissione della Comunità nel 1989, e la proposta per disciplinare la responsabilità civile per danni derivanti da attività pericolose per l'ambiente[46], presentata dal Consiglio d'Europa nel 1992, prevedono entrambe l'obbligo della copertura assicurativa, oppure di garanzie finanziarie. Più cauto appare il Libro Verde del 1993 che esamina estesamente il tema della riparazione dei danni causati all'ambiente, mettendo in luce pregi e debolezze dell'assicurazione obbligatoria: "L'assicurazione funge da importante meccanismo di indennizzo.... può costituire un deterrente in quanto promuove una migliore prevenzione degli incidenti". Tuttavia, "le incertezze connesse alla responsabilità civile nell'ambito del danno ambientale sollevano numerosi problemi anche per l'assicurazione. Gli assicuratori esitano nel garantire la copertura se sono incerti circa i tipi e le probabilità del danno che può avere luogo o se perdite imprevedibili riducono il fondo costituito... spesso essi reagiscono aumentando i prezzi dei premi o ritirandosi dal mercato dell'assicurazione di responsabilità civile connessa a danni ambientali... ". Comunque, "la valutazione del rischio operata dall'industria assicurativa è benefica in quanto, mentre riduce il rischio di una perdita economica per l'assicuratore, riduce anche il rischio di danno ambientale". Tale posizione cauta e attenta della Commissione riflette la complessità dell'argomento e incorpora, in un certo senso, le reazioni delle parti coinvolte. Gli assicuratori temono, da una parte indennizzi elevati, dall'altra il vincolo creato dall'obbligatorietà. L'analisi tecnica preliminare al contratto di assicurazione, infatti, potrebbe essere destituita della sua utilità perché, essendo la compagnia di assicurazione obbligata a concedere all'impresa la copertura assicurativa, essa non sarebbe più libera di far dipendere la copertura dall'adozione di misure preventive da parte dell'impresa. A loro volta le imprese, essendo obbligate ad avere dei rapporti con la compagnia di assicurazione, si preoccupano della possibilità che questa imponga premi elevati e che condizioni le loro strategie di protezione ambientale. In sintesi, il quadro è estremamente

46 "Draft convention on civil liability for damage resulting from activities dangerous to the environment", Strasburgo 1992.

complesso ed appare difficile individuare una conclusione univoca, di accettazione o di rifiuto dello strumento.

- I fondi di indennizzo costituiscono uno strumento di copertura del danno che integra l'assicurazione in quei casi nei quali le dimensioni del danno eccedano la copertura assicurativa, oppure non sia possibile identificare i responsabili. Essi possono essere privati o pubblici, volontari o resi obbligatori dalla legge. Il loro finanziamento può essere costituito da versamenti delle imprese oppure da contributi pubblici oppure da imposte specifiche per il settore nel quale esiste il fondo. La loro finalità, essenzialmente, è quella di essere di ausilio allo strumento dell'assicurazione nella sua funzione di indennizzo delle vittime, garantendo compensazioni più veloci e più sicure. Esempi di fondi di indennizzo esistono nei singoli paesi. In Olanda, istituito nel 1972, vi è un fondo per la compensazione dei danni procurati dall'inquinamento atmosferico; in Francia, ve ne è uno per l'inquinamento acustico procurato dagli aerei, esso viene finanziato dalle compagnie di assicurazione ed è a beneficio delle persone che abitano in prossimità degli aeroporti; in Svezia, esiste un fondo finanziato dalle imprese e del quale beneficiano le vittime nei casi in cui l'autore del danno sia ignoto oppure insolvente oppure non dichiarabile responsabile dalla legge. Infine, naturalmente, il Superfund USA, finanziato per mezzo di tasse dalle imprese petrolifere e chimiche statunitensi. L'industria del petrolio, in particolare, offre numerosi esempi di fondi di indennizzo. Nel 1969, ad esempio, la Convenzione di Bruxelles ha definito i limiti di responsabilità del proprietario della nave che, in caso di oil spill, è responsabile primario del danno ed è tenuto a risarcire le vittime. Qualora, però, sulla base delle clausole di esonero, il proprietario della nave non sia responsabile, le vittime vengono risarcite da un fondo, attivo dal 1971, denominato Fondo internazionale di indennizzo.[47] Altri esempi di fondi di indennizzo nell'industria petrolifera sono:

- OPOL (Offshore Pollution Liability Agreement), operante dal 1974: indennizza danni da esplorazione petrolifera e sfruttamento della piattaforma continentale.

[47] Per una panoramica sulla responsabilità per i danni da idrocarburi, Cassar, G. 1993.

- TOVALOP (Tankers Owners Voluntary Agreement Concerning Liability for Oil Pollution), operante dal 1978: rimborsa le spese governative per la protezione e la bonifica delle coste.

- CRISTAL (Contract Regarding Interim Supplement to Tanker Liability for Oil Pollution), operante dal 1978: compensa le vittime per la quota eccedente il limite di garanzia del Tovalop.

  Infine, occorre ricordare l'orientamento della Comunità Europea che ha proposto l'istituzione di un fondo di indennizzo per danni causati da rifiuti.

Naturalmente, un'investigazione più ravvicinata del mondo dell'assicurazione ambientale rileverebbe l'esistenza di una pluralità di strumenti e di tendenze qui non segnalate. In Olanda, Francia e Italia, ad esempio, alcune compagnie di assicurazione si sono raggruppate in pool per far fronte a danni di rilevanza tale da non poter essere sopportati da un solo assicuratore. Altra possibilità è costituita dai captive, organismi fondati dalle imprese per ripartire meglio il rischio e per gestirlo in modo più efficiente. In sintesi, l'adattamento dei contratti di assicurazione alle specificità del bene ambiente, così come la creazione di istituzioni in grado di gestire in modo efficiente le sue tipicità, rappresentano un fenomeno in divenire, ancora suscettibile di affinamenti e progressi.

4.3. Il processo.

Gli agenti ricorrono al giudizio imparziale di una Corte di Giustizia quando, una volta verificatosi il danno, non riescono ad accordarsi circa la compensazione. Idealmente, la Corte alloca la responsabilità tra gli agenti in modo esatto: in caso di responsabilità per colpa, si atterrà alla regola del giudice Hand, in caso di responsabilità oggettiva, rileverà se vi è causalità

tra azione dell'injurer e danno e deciderà di conseguenza. La compensazione stabilita sarà pari al valore del danno. Tuttavia, la Corte spesso non dispone di perfetta informazione sicché il suo giudizio può allontanarsi dallo schema ideale. D'altra parte, in un mondo di perfetta informazione gli agenti sarebbero in grado di prevedere il giudizio della Corte e, quindi, lo troverebbero superfluo. Pertanto, agenti razionali ricorrono al giudice essenzialmente perché vi è incertezza e sono possibili previsioni diverse circa l'esito del processo. La fiducia nella vittoria indurrà un agente a rivolgersi alla Corte; l'alta probabilità di una sconfitta potrà distoglierlo da questa prospettiva. Inoltre, analogamente a quanto accade nella contrattazione privata, le spese giudiziali ed il modo in cui esse sono ripartite tra gli agenti condizionano i comportamenti. Nel sistema inglese le spese del processo sono a carico di chi perde, mentre in quello americano ciascun agente (accusatore e accusato) paga le sue spese. Pertanto, nel sistema inglese una sconfitta processuale è molto più onerosa che in quello americano e il comportamento degli agenti terrà conto di tale maggiore onerosità.

Sinteticamente, gli elementi che influiscono sul comportamento di un agente (accusatore che ha subito il danno), inducendolo all'accordo oppure al ricorso al giudice, sono:

- la vincita attesa (V) che dipenderà dal valore del danno subito (D) e dalla probabilità di vittoria (P, che a sua volta dipenderà dal pessimismo o dall'ottimismo dell'agente circa l'esito del processo). Si avrà, quindi: $V = PD$.

- le spese giudiziali attese di chi accusa (C1) e di chi si difende (C2). Poiché con la regola americana gli agenti sostengono le proprie spese giudiziali sia in caso di vittoria che di sconfitta, esse devono essere comunque sottratte alla vincita attesa. Pertanto, la vincita attesa netta (VN) dell'accusatore è: $PD - C1$. Nel caso inglese, invece, l'accusatore dovrà tenere conto anche delle spese giudiziali di chi si difende (C2) perché, in caso di sconfitta, egli sarà tenuto a pagarle. Poiché la probabilità di sconfitta è pari a $1 - P$, la vincita attesa netta sarà: $VN = PD - (1 - P)(C1 + C2)$.

- l'attitudine rispetto al rischio che, in caso di avversione (propensione), indurrà l'agente a sottrarre dalla (aggiungere alla) sua vincita attesa netta un premio di rischio. In caso di neutralità al rischio, invece, l'agente considererà nel suo calcolo economico il valore della vincita attesa netta senza né diminuirlo né aumentarlo.

Dati tali elementi fondamentali, l'incentivo di un agente a ricorrere in giudizio crescerà, ceteris paribus:

- al crescere della vincita attesa, ovvero al crescere del suo ottimismo circa l'ipotesi di una sua vittoria (aumento della probabilità di vittoria).
- al diminuire delle spese giudiziali attese. Sistema legali efficienti, cause brevi, basse spese giudiziali avvicinano gli agenti alla Corte di Giustizia, e viceversa.
- al crescere della sua propensione al rischio.

Al contrario, un agente troverà più conveniente l'accordo privato, ovvero il suo incentivo a ricorrere al giudice diminuirà, al diminuire della vincita attesa (diminuzione della probabilità di vittoria), all'aumentare delle spese giudiziali, al crescere della sua avversione al rischio.

Possiamo, a questo punto, pervenire ad alcune conclusioni circa la convenienza per gli agenti di un accordo privato oppure del ricorso al giudice. Per semplicità, consideriamo l'ipotesi che entrambi gli agenti (accusatore e imputato) siano neutrali rispetto al rischio e che assegnino la stessa probabilità alla possibilità di vittoria dell'accusatore (ovvero P è identico per entrambi). Nel sistema americano, la vincita attesa netta dell'accusatore sarà PD - C1, mentre il pagamento atteso, includendo le spese giudiziali, è per chi si difende: PD + C2. Pertanto, l'accusatore troverà conveniente qualsiasi accordo che lo compensi in misura uguale o maggiore di PD - C1, mentre chi si difende troverà conveniente un accordo che comporti una compensazione minore o uguale a PD + C2. In altre parole, vi sarà un accordo se la compensazione cade nell'intervallo compreso tra PD - C1 e PD + C2, estremi inclusi. E ciò significa, poiché le spese giudiziali C1 e C2 rispettivamente abbassano ed alzano l'estremo inferiore e l'estremo superiore dell'intervallo di accordo, che tale intervallo crescerà al crescere

dei costi del processo: più precisamente, l'intervallo di accordo sarà pari alla somma delle spese giudiziali C1 e C2. Nel caso del sistema inglese, invece, saranno convenienti accordi che diano luogo a compensazioni comprese tra PD - (1 - P)(C1 + C2), la vincita attesa netta di chi accusa, e PD + P(C1 + C2), che rappresenta il pagamento atteso, comprensivo delle spese giudiziali, di chi si difende. Di nuovo, l'intervallo di accordo sarà pari alla somma delle spese giudiziali C1 e C2. Questo risultato è d'altra parte intuitivo poiché afferma che una contrattazione privata tra le parti sarà tanto più conveniente quanto più costoso è il ricorso alla giustizia.

Naturalmente, l'ipotesi che la probabilità di vittoria di chi accusa (P) sia uguale per entrambi gli agenti può essere rimossa. E' intuitivo e facilmente dimostrabile che l'intervallo di accordo cresce, ceteris paribus, al crescere del pessimismo dell'agente che potrebbe ricorrere in giudizio. Mentre, al contrario, l'aumentare del suo ottimismo riduce il range di accordo e, in alcuni casi, può farlo scomparire. In altre parole, ci sono situazioni nelle quali colui che ricorre in giudizio è così sicuro della propria vittoria che la sua vincita attesa è maggiore del pagamento atteso dell'agente che si difende e, pertanto, nessun accordo è possibile ed il ricorso al giudice costituisce l'unica soluzione. Analogamente, ceteris paribus, l'avversione al rischio degli agenti aumenta l'intervallo di accordo mentre la propensione al rischio lo riduce, finché, in casi estremi, esso scompare.

Questi elementi - incertezza sull'esito della causa, spese giudiziali, attitudine al rischio - rappresentano le forze principali coinvolte in un processo. Naturalmente, attraverso la considerazione di elementi quali, ad esempio, la *reputation*, il livello di informazione degli agenti, il loro comportamento strategico, è possibile accrescere il grado di realismo e di complessità della spiegazione e pervenire a risultati più completi.

5. CONCLUSIONI

E' utile, al termine del lavoro, ripercorrere, in sintesi, le principali conclusioni.

- Le interazioni tra agenti generate dall'insorgere di un'esternalità negativa sono state esaminate nell'ambito di un percorso ai cui estremi vi sono, da una parte, l'accordo, dall'altra, il ricorso ad una Corte di Giustizia. Si è visto che la possibilità di un accordo è ostacolata dai costi di transazione che possono in alcuni casi, quando sono molto alti, rendere l'accordo non conveniente. Inoltre, qualora si raggiungesse un accordo, esso sarebbe influenzato dall'esistenza dei costi di transazione. In particolare, contrariamente alla tipica contrattazione alla Coase, non è più vero che il punto di equilibrio in cui si raggiunge l'accordo è indipendente da chi detiene i diritti di proprietà. Al contrario, la detenzione dei diritti di proprietà assicura all'agente un vantaggio distributivo. Ciò significa che lo status quo, ovvero il contesto istituzionale in cui si svolge la contrattazione, ne influenza l'esito: sono possibili più punti di equilibrio e, dunque, più livelli di inquinamento "ottimo". Poiché livelli diversi di inquinamento "fisico" possono essere ritenuti ottimi appare importante riflettere sul significato dell'espressione "livello ottimo di inquinamento" e sull'opportunità di considerare lo status quo un elemento da sottoporre a giudizio oppure da prendere come dato ingiudicabile.

- Un diritto può essere protetto da ingiunzione oppure da compensazione. Il modo in cui i diritti sono protetti influenza l'esito della contrattazione qualora si ammetta la possibilità di comportamenti strategici. Vi è spazio per un comportamento strategico, e per soluzioni Pareto-inefficienti, se il danneggiato possiede un diritto assoluto di ingiunzione oppure se

egli è protetto da compensazione ma le autorità, a causa di un deficit informativo, fissano un livello di compensazione maggiore del danno reale. Più in generale, se le autorità non dispongono di informazione perfetta vi è possibilità di comportamenti strategici e di risultati inefficienti, sia in caso di ingiunzione che di compensazione.

- Si distinguono due regimi di responsabilità: responsabilità per colpa e responsabilità oggettiva. In caso di esternalità negativa, nel primo regime, la responsabilità è assegnata sulla base di un giudizio sulla diligenza dell'agente nello svolgimento della propria attività; nella seconda, l'agente è sempre ritenuto responsabile dell'incidente che ha causato, indipendentemente dal suo grado di diligenza. Nonostante siano caratterizzati da due principi diversi, questi due regimi attivano il medesimo livello di prevenzione. Differenze esistono per quanto concerne gli effetti sul livello di attività e su prezzi e consumi: solo quando vige un regime di responsabilità oggettiva l'inquinatore è indotto a moderare il suo livello di attività ed i prezzi riflettono l'intero costo sociale del danno e, dunque, il livello dei consumi e dell'output è efficiente. Inoltre, a differenza della responsabilità per colpa, la responsabilità oggettiva, "assicura" sempre la vittima contro il rischio di incidente, è caratterizzata da un più basso costo unitario delle cause e, quindi, da un più alto numero di cause.

- Da un confronto con altri strumenti della politica ambientale emerge che la responsabilità oggettiva sta alla tassa pigouviana come la responsabilità per colpa sta allo standard. Tuttavia, differenze esistono per quanto riguarda sia l'entità dei pagamenti effettuati dai danneggianti sia il loro destinatario. Inoltre, a differenza della tassa e dello standard, i regimi di responsabilità: non richiedono informazione perfetta da parte delle autorità sulla funzione di danno e, quindi, spostano il costo dell'informazione dal settore pubblico a quello privato; non richiedono sistemi di monitoraggio permanenti né meccanismi di raccolta e gestione delle compensazioni, Tuttavia, essi sono strumenti poco efficienti in caso di inquinamento continuo e possono appesantire l'apparato giudiziario comportando costi di gestione molto elevati.

- L'applicazione dei regimi di responsabilità all'ambiente può rivelarsi problematica a ragione di difficoltà che sorgono nell'individuazione dei danni compensabili e del nesso di causalità tra azione e incidente. Pluralità di cause, danni sinergici, ritardi temporali sono alcuni elementi che possono ostacolare l'individuazione del nesso di causalità. Infine, i tradizionali ostacoli che esistono nel valutare il danno rafforzano le difficoltà di applicazione dei regimi di responsabilità all'ambiente.

- Tra le tendenze recenti, particolarmente interessanti appaiono le indicazioni contenute nel "Libro verde sul risarcimento dei danni all'ambiente" della Comunità Europea ed il CERCLA (Comprehensive Environmental Response, Compensation and Liability Act) statunitense. Il primo documento costituisce una riflessione sistematica sull'impiego dello strumento della responsabilità civile per i casi di danni ambientali. Esso individua le caratteristiche della responsabilità civile applicata all'ambiente e, con moderazione, suggerisce l'adozione del regime della responsabilità oggettiva. Il CERCLA statunitense, invece, rappresenta un punto di riferimento della legislazione ambientale mondiale. In esso, l'ambiente viene riconosciuto come bene giuridico autonomo, si adotta la responsabilità oggettiva, si dispone un meccanismo di azione rapida di pulizia e di risanamento di siti contaminati.

- Gli agenti possono coprire il rischio di incidenti ambientali assicurandosi e trasferendo, così, il rischio sulla compagnia di assicurazione. Tale copertura del rischio svolge un'importante funzione perché assicura il danneggiato contro possibili insolvenze dell'injurer (giustizia sociale) ed incentiva la prevenzione poiché gli assicuratori, adeguando i premi alla dose di rischio dei potenziali danneggianti, li inducono ad una più attenta prevenzione (efficienza). Tuttavia, le tipicità di fenomeni ambientali - incertezza nel nesso di causalità, nel fenomeno inquinamento e nella valutazione del danno -, rafforzati dai tradizionali problemi di moral hazard e adverse selection, ostacolano lo strumento assicurativo indebolendone l'efficacia.

- Tra le recenti esperienze di assicurazione contro il rischio ambientale, appaiono interessanti lo strumento dell'assicurazione obbligatoria ed i fondi di indennizzo. La prima è stata introdotta in Germania e, con moderazione, sostenuta in un paio di documenti dalla Comunità Europea che però, nel "Libro verde" del 1993 ne evidenzia anche i limiti. Il secondo strumento, invece, costituisce un'integrazione dell'assicurazione in casi in cui le dimensioni del danno eccedano la copertura assicurativa oppure non sia possibile identificare i responsabili. Numerose sono state, nell'ultimo ventennio, le esperienze di fondi di indennizzo, in particolare nell'industria del petrolio.

- Quando l'accordo non è possibile gli agenti ricorrono al giudizio di una Corte di Giustizia. Idealmente, la Corte alloca la responsabilità tra gli agenti in modo oggettivo, ad esempio, nel caso in cui sia vigente un regime di responsabilità per colpa, attenendosi alla regola di Hand. La decisione degli agenti - se accordarsi oppure ricorrere al giudice - si fonda su un calcolo razionale in cui le variabili principali prese in considerazione sono la vincita attesa e le spese giudiziali. Ottimismo e pessimismo degli agenti, regole diverse di ripartizione delle spese processuali, avversione e propensione al rischio costituiscono altri elementi che influenzano il comportamento degli agenti.

In sintesi, queste sono le principali conclusioni del lavoro. In esso si è cercato di documentare le interrelazioni tra analisi economica del diritto e ambiente ed il ruolo di alcuni strumenti particolarmente rilevanti. Ancora una volta, ad ogni modo, appare opportuno ricordare l'inevitabile dose di parzialità e incompletezza intrinseca in ogni conclusione che, come tutto ciò che è rappresentazione - tenue telaio - del reale, poggia necessariamente su un insieme di assunzioni che, implicite o esplicite, ne limitano la robustezza. Al di là della rappresentazione, imprendibile, sta il reale - un tessuto ben più fitto del telaio che lo sottende.

BIBLIOGRAFIA

Alpa, G. 1990, Danni catastrofici e limiti alla responsabilità d'impresa, *L'Impresa Ambiente*, 4/90, pp. 119 - 120.

Arrow, Kenneth et al. 1993, *Report of the NOAA Panel on Contingent Valuation*, Washington D.C.

Ball, S. and Stuart, B. 1991, *Environmental Law*, Blackstone Press Limited, London.

Brealey, M. 1993 (edited by), *Environmental Liabilities and Regulation in Europe*, The Hague: International Businesses Publishing.

Bromley, D. W. 1991, *Environment and Economy. Property Rights and Public Policy*, Blackwell, Oxford.

Calabresi, G. 1968, Transaction Costs, Resource Allocation, and Liability Rules in, Dorfman, R. and Dorfman, N. S. (edited by), 1977 (ripreso da: *The Journal of Law and Economics*, April 1968).

Capria, A. 1990, Si chiama Superfund il risanamento made in USA, *L'Impresa Ambiente*, 4/90, pp. 114 - 118.

Cassar, G. 1993, Acque difficili per i danni da idrocarburi, *L'Impresa Ambiente*, 10/93, pp. 68 - 71.

Cellerino, R. 1993, *Oltre la tassazione ambientale. Nuovi strumenti per il controllo dell'inquinamento*, Il Mulino, Bologna.

Colacicchi, M. 1993, Dal mondo assicurativo un'opinione sul rischio ambientale, *L'Impresa Ambiente*, 10/93, pp. 72 - 73.

Chance, C. (prepared by) 1991, *Environmental Law Guide. An Industry Report*, LLoyd's of London Press Ltd, London.

Commissione delle Comunità Europee 1993, *Il libro verde sul risarcimento dei danni all'ambiente*, 14/5/93, Bruxelles.

Coase, R. H. 1990, *The Firm the Market and the Law*, The University of Chicago Press, Chicago.

Cooter, R. D. 1991 and Rubenfield Daniel L., Economic Theories of Legal Liabilities, *The Journal of Economic Perspectives*, Vol. 5, N. 3.

Dales, J. H. 1968, The Property Interface, in Dorfman, Robert and Dorfman, Nancy S. (edited by) 1977.

De Strobel, D. 1990, Problemi di assicurazione del danno ambientale, *L'Impresa Ambiente*, 4/90, pp. 109 - 113.

De Strobel, D. 1992, Dalle compagnie assicurative gli strumenti per la prevenzione, *L'Impresa Ambiente*, 3/92, pp. 20 - 26.

Di Giovanni, B. 1995, *Civil liability for environmental damage in the business community: a European comparative study,* unpublished EAEME dissertation.

Dorfman, R. and Dorfman, N. S. (edited by) 1977, *Economics of the Environment. Selected Readings*, Northon & Company, New York.

Ferraiolo, D. 1992, La sindrome emergenza è il rischio più grave, *L'Impresa Ambiente*, 3/92, pp. 4 - 7.

Fenn, P. and McGuire, A. 1994, The Assessment: The Economics of Legal Reform, in Fenn, P. et al. 1994, Economics of Legal Reform, *Oxford Review of Economic Policy*, Vol. 10, No. 1.

Freeman III, M. A. 1993, *The Measurement of Environmental and Resource Values. Theory and Methods*, Resources for the Future, Washington, D. C..

Hughes, J. W. and Snyder, E. A. 1995, Litigation and settlement under the English and American rules: theory and evidence, *The Journal of Law and Economics*, April 1995.

IBA (International Bar Association), 1990, *Environmental Liability. 7th Residential Seminar on Environmental Law*, Graham & Trotman and IBA.

Mattei, U. e Pulitini, F. (a cura di) 1994, *Consumatore Ambiente Concorrenza. Analisi economica del diritto*, Giuffrè, Milano.

Misch, A. 1994, *Valutare i rischi ambientali per la salute*, in Lester R. Brown e altri, *State of the World 1994. Rapporto sul nostro pianeta del Worldwatch Institute*, Isedi, Torino.

Milgrom, P. and Roberts, J. 1992, *Economics, Organization and Management*, Prentice Hall International,, Inc., London.

Mulè, G. 1992, Nell'industria di processo nuovi spazi per l'assicurazione, *L'Impresa Ambiente*, 3/92, pp. 31 - 35.

OECD 1989, *Economic instruments for environmental protection*, OECD, Paris.

Panella G. 1993, "Esternalità, diritti di proprietà e responsabilità civile", in Gerelli E. (a cura di), "Mercato unico e ambiente - Contrasto o compatibilità?", Pirola, Milano.

Pavesi, F. 1992, Ma la polizza verde non entusiasma le imprese, *L'Impresa Ambiente*, 3/92, pp. 27 - 29.

Pesaro, G. 1994, Le implicazione economico-giuridiche delle normative europee sul danno ambientale, *Rivista delle fonti di energia e dell'ambiente,* n. 1, pp. 149 - 164.

Pesaro, G. 1994, L'impresa e la responsabilità per danno ambientale, *Rivista delle fonti di energia e dell'ambiente,* n. 1, pp. 165 - 189.

Polinsky, Mitchell A. 1980, Strict Liability vs. Negligence in a Market Setting, *American Economic Review, Papers & Proceedings*, vol. 70 no. 2, pp. 363 - 367.

Polinsky, Mitchell A. 1989, *An introduction to law and economics*, Little, Brown and Company, Boston.

Posner, Richard A. 1992, *Economic Analysis of Law*, Little Brown and Company, Boston.

Pozzo, B. 1994, Il criterio di imputazione della responsabilità per danno all'ambiente nelle recenti leggi ecologiche, in Trimarchi, Pietro (a cura di) 1994.

Rapisarda Sassoon, C. 1990, Il danno ambientale tra colpa e responsabilità oggettiva, *L'Impresa Ambiente*, 4/90, pp. 99 - 104.

Rapisarda Sassoon, C. 1994 (a cura di), *Capire l'ambiente. Guida tecnico-normativa ai controlli ambientali,* Il Sole 24 Ore Libri, Milano.

Sagoff, Mark 1990, *The economy of the earth. Philosophy, law and the environment*, Cambridge University Press, Cambridge.

Sassoon, E. e Rapisarda Sassoon, C. 1993, *Management dell'ambiente*, Il Sole 24 Ore Libri, Milano.

Trimarchi, Pietro (a cura di) 1994, *Per una riforma della responsabilità civile per danno all'ambiente*, Giuffrè Editore, Milano.

Villa G., 1994, Nesso di causalità e responsabilità civile per danni all'ambiente, in Trimarchi P. (a cura di), 1994.

Williamson, O. E. 1987, *Le istituzioni economiche del capitalismo. Imprese, mercati, rapporti contrattuali*, Franco Angeli, Milano.

Yeager, P. C. 1991, *The Limits of Law. The Public Regulation of Private Pollution*, Cambridge University Press, Cambridge.

Zweifel, P. and Tyran J.-R. 1994, Environmental impairment liability as an instrument of environmental policy, *Ecological Economics 11*, pp. 43 - 56.

Économie publique/Public economics

07 | 2001/1
Analyse économique du droit

Economic Analysis of environmental Law : An Introduction

Michael G. Faure

Édition électronique
URL : http://journals.openedition.org/economiepublique/1592
DOI : 10.4000/economiepublique.1592
ISSN : 1778-7440

Éditeur
IDEP - Institut d'économie publique

Édition imprimée
Date de publication : 15 janvier 2001
ISBN : 2-8041-3634-5
ISSN : 1373-8496

Référence électronique
Michael G. Faure, « Economic Analysis of environmental Law : An Introduction », *Économie publique/ Public economics* [En ligne], 07 | 2001/1, mis en ligne le 07 décembre 2005, consulté le 12 septembre 2020. URL : http://journals.openedition.org/economiepublique/1592 ; DOI : https://doi.org/10.4000/economiepublique.1592

© Tous droits réservés

économie publique

public economics

Revue de **l'Institut d'Économie Publique**

Deux numéros par an

nº 7 – 2001/1

IDEP

© De Boeck & Larcier s.a., 2001
Editions De Boeck Université
Rue des Minimes 39, B-1000 Bruxelles

Tous droits réservés pour tous pays.
Il est interdit, sauf accord préalable et écrit de l'éditeur,de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

Imprimé en Belgique

ISSN 1373-8496
Dépôt légal 2001/0074/226
ISBN 2-8041-3634-5

économie publique sur internet : www.economie-publique.fr

© Institut d'économie publique – IDEP

Centre de la Vieille-Charité

2, rue de la Charité – F-13002 Marseille

Tous droits réservés pour tous pays.

Il est interdit, sauf accord préalable et écrit de l'éditeur, de reproduire (notamment par photocopie) partiellement ou totalement le présent ouvrage, de le stocker dans une banque de données ou de le communiquer au public, sous quelque forme et de quelque manière que ce soit.

La revue économie publique bénéficie du soutien du Conseil régional Provence-Alpes-Côte d'Azur

ISSN 1373-8496

Economic Analysis of environmental Law : An Introduction

Michael G. Faure
Université de Maastricht

1 Introduction

The problem of environmental pollution is caused by the fact that some activities which are as such beneficial, such as the production of pharmaceuticals, can cause negative side effects for third parties, the traditional externalities. Precisely because pollution is an externality, the starting point of the economic analysis of environmental pollution is that a decision maker, such as the pharmaceutical company of our example, will not take into account the externality when it takes decisions on e.g. the production level and the investments in measures to avoid pollution, such as the instalment of a water treatment plant. Environmental pollution is considered by many scholars to be the example of an externality (see generally Van den Bergh, 1988, 234).

In the absence of the law there will – in principle - be no incentive for the polluting factory to take in account the pollution it is causing. In other words : in the absence of legal rule the externality will not be internalised. This immediately indicates in a very simple way the economic goal of environmental law : it should lead to an internalisation of the externality by forcing the potential polluter to take into account the pollution it is causing in its decision making process (Faure, 1996). If the law can reach this, the pollution would no longer be external to the activity, but would be internalised e.g. because the potential polluter decides as a consequence of the pollution to invest in abatement techniques. The body of environmental law subsequently deals with the question how environmental law can give incentives to internalise the externality the pollution is causing. A variety of legal instruments can lead to such an internalisation. Economists would traditionally advance the use of taxes to reach this goal, although increasingly attention is paid to market based instruments, such as emissions trading and marketable pollution rights.

In this paper I will try to show how a variety of traditional legal

instruments can be used to remedy environmental pollution. First the crucial question will be addressed whether liability can be consi-dered as an instrument to prevent environmental pollution. This is an important question, since many studies on environmental economics seem to neglect the importance of environmental liability. Liability rules are considered by some also as a market solution and therefore certainly merit attention. However, the traditional public interest criteria for regulation can be used to explain why liability rules alone can not suffice to control the risks posed by environmental harm (2). However, the conclusion that in environmental law some regulatory invention (through licences and emission standards) will be necessary does not exclude the role of environmental liability. Hence the question arises how liability rules and regulation can be used jointly to remedy environmental pollution and how they mutually influence each other (3). Finally, the question will have to be asked which of these liability rules is optimal to internalize environmental risks (4) and a few concluding remarks will be formulated.

Thus this paper on the one hand attempts to show, without fancy economic modelling, how traditional economic analyses (theory of regulation, liability rules) can be used to analyse environmental law. This may provide useful insights to lawyers wondering about the potential use and applications of "law and economics" to environmental law. The concrete application of well known theories to the specific field of environmental law has the advantage that it immediately shows the usefulness of economic theory for legal doctrine and policy. At the same time, a reminder of the classic theories of law and economics may be useful for environmental economists as well who on the one hand some times tend to be overenthousiastic of market based instruments, forgetting some benefits of regulation (mainly information advantages) and sometimes neglecting that liability rules can be considered as a market solution as well.

Turning back to the crucial question of this paper, how the law can contribute to an internalisation of environmental harm, the first crucial question is obviously whether environmental risks should primarily be internalized via regulation or liability rules.

2 Environmental regulation

2.1 Variety of legal instruments

Assuming that, as will be the case in many pollution cases, Coasean

bargaining is not possible because of prohibitive transaction costs[1], the question arises what kind of policy instruments should be used to give incentives to a potential polluter to prevent environmental harm. Traditionally there were three possible instruments which were addressed. First of all it is possible to tax the pollution and thus to use a system of levies or charges, which will give the potential polluter an incentive to reduce environmental harm. Second, it is possible to use the liability system, assuming that the potential polluter will be deterred by the foresight of having to pay compensation to a victim for the environmental harm he caused. Third, it is possible to fix pollution standards (notably emmission standards) *ex ante* in regulation and more specifically in environmental licenses. Now, in addition to these, a whole new set of policy instruments has been developed. Economists increasingly advocate the use of market oriented policy instruments, such as systems of emmission trading and marketable permits. In addition attention is given to voluntary compliance mechanisms, such as environmental agreements.

It is obviously not possible to discuss this whole set of possible environmental policy instruments within the scope of this paper[2]. I will focus on two traditional instruments, being on the one hand liability rules and on the other hand safety regulation.

2.2 Criteria for safety regulation

Let us examine under what kind of circumstances liability rules may not suffice to deter environmental harm and a regulatory intervention may be necessary. The choice between regulation and liability rules has been thoroughly examined by Steven Shavell in 1984, in a paper in which he advances several criteria that influence the choice between safety regulation and liability rules (see Shavell, 1984 and 1987).

2.2.1 Information asymmetry as a criterion for regulatory intervention

Information deficiencies have often been advanced as a cause of market failure and as the justification for government intervention through regulation (see Stigler, 1961; Schwartz and Wilde, 1979; Mackaay, 1982). Also, for the proper operation of a liability system, information on e.g. the existing legal rules, the accident risk, and efficient measures to prevent accidents, is a precondition for an efficient deterrence. According to Shavell, the parties in an accident setting generally have much better information on the accident risk than that possessed by the regulatory body (see Shavell, 1984, 359). The parties themselves have, in

[1] We assume here indeed positive transactions costs so that this internalisation can not automatically be reached via the Coase Theorem (see Coase, 1960).

[2] An excellent overview of these instruments is presented by Gunningham and Grabosky, 1998.

principle, the best information on the costs and benefits of the activity that they undertake and of the optimal way to prevent accidents. This "assumption of information" will, however, be reversed if it becomes clear that some risks are not readily appreciated by the parties in an accident setting. Therefore, for every activity the question that will have to be asked is whether either the government or the parties involved can acquire the information at the least cost.

2.2.2 Insolvency risk

If the potential damages can be so high that they will exceed the wealth of the individual injurer, liability rules will not provide optimal incentives. The reason is that the costs of care are directly related to the magnitude of the expected damages. If the expected damages are much greater than the individual wealth of the injurer, the injurer will only consider the accident as having a magnitude equal to his wealth. He will take, therefore, only the care necessary to avoid an accident equal to his wealth, which can be lower than the care required to avoid the total accident risk (see Shavell, 1984, 360). This is a simple application of the principle that the deterrent effect of tort liability only works if the injurer has assets to pay for the damages he causes. If an injurer is protected against such liability, a problem of underdeterrence arises[3] (see Shavell, 1986).

Safety regulation can overcome this problem of underdeterrence caused by insolvency. [4]In that case the efficient care will be determined *ex ante* by regulation and will be effected by enforcement instruments which induce the potential injurer to comply with the regulatory standard, irrespective of his wealth.

In that case a problem might still arise if the regulation were also enforced by means of monetary sanctions. Again, if these were to exceed the injurer's wealth, the insolvency problem would remain. Hence, if a safety regulation is introduced because of a potential insolvency problem, the regulation itself should be enforced by non-monetary sanctions (see Shavell, 1985).

2.2.3 The threat of a liability suit

Some activities can cause considerable damage, but even so a law suit to recover these damages may be never brought. If this were the case, there would of course be no deterrent effect of liability rules. Therefore, the absence of a liability suit would again be an argument to enforce the duty of efficient care by means of safety regulations rather than through liability rules (see Shavell, 1984, 363). There can be a number

[3] Below we will show that insolvency causes especially a problem under a strict liability rule, but less so under negligence.

[4] If insurance would come into the picture it could overcome the problems of underdeterrence, provided that the moral hazard problem, caused by insurance, can be cured.

of reasons why a law suit is not brought, even though considerable damages have been caused.

Sometimes an injurer can escape liability because the harm is thinly spread among a number of victims. As a consequence, the damage incurred by every individual victim is so small that he has no incentive to bring a suit. In particular, this problem will arise if the damage is not caused to an individual but to a common property, such as e.g. the surface waters in which each member of the population has a minor interest. In addition, a long time might have elapsed before the damage becomes apparent; in this case much of the necessary evidence may be either lost or not obtained. Another problem is that if the damage only manifests itself years after the activity, the injurer might have gone out of business.

A related problem is that it is often hard to prove that a causal link exists between an activity and a type of damage (see Landes and Posner, 1984, 417). The burden of proof of a causal relationship becomes more difficult with the increasing passage of time since the damaging incident took place. Often a victim will not recognise that the harm had been caused by a tort, but might think that his particular ailment, e.g. cancer, had a "natural cause", associated with a general ill health . For all these reasons a liability suit might not be brought and hence safety regulation is necessary to ensure that the potential polluter takes efficient care[5].

2.2.4 Administrative costs

When examining the pro and contra's of liability versus regulation, the administrative costs of both systems should also be compared. Liability rules are clearly costly in terms of time for both parties and in court fees. A part of these costs is borne by the whole community, such as e.g. the cost of the legal system, fees for the judges etc. Regulation produces costs for the community, including the costs of making the regulation, setting the standards, passing the statutes etc. and of subsequent enforcement (see Shavell, 1984, 363-364).

In this respect the liability system seems to have an advantage : the administrative costs of the court system are only incurred if an accident has actually happened. The main advantage of the tort system is that a lot of accidents will be prevented by the deterrent effect of being held liable and having to pay damages to the victim. In case of safety regulation the costs of passing the regulation and of enforcing it are always there, whether there are accidents or not.

[5] For alternatives to liability suits see Bocken, 1987, 83-87; 1988, 3-10.

2.3 The need to regulate environmental pollution

After having discussed these criteria for regulation[6] I will now discuss the question of how these criteria relate to environmental pollution. If one takes the criteria for safety regulation discussed above and applies them to the potential risk caused by environmental pollution, there is no doubt that liability rules alone are not sufficient.

If one looks at the first criterion, that of information costs, it must be stressed that an assessment of the risks of a certain activity often requires expert knowledge and judgement. Small organisations might lack the incentive or resources to invest in research to find out what the optimal care level would be. Also, there would be little incentive to carry out intensive research if the results were automatically available to competitors in the market : this is the well-known "free rider" problem. This problem can partially be countered by legal instruments granting an intellectual property to the results of the research. However, the problem remains that it may not be possible for small companies to undertake studies on the optimal technology for preventing environmental damage. Therefore, it is often more efficient to allow the government itself to do the research on the optimal technology (e.g. in a governmental environmental research institute). The results of this research can then be passed on to the parties in the market through the regulation. Hence, the setting of environmental standards in regulation can be seen as a means of passing on information on the minimal environmental technology required. Obviously, it is more efficient for the government to acquire information on the optimal emission standard than it would be for e.g. an individual firm to find out what additional reduction in pollution would produce an optimal reduction of the expected damages from the emission. There are undeniable "economies of scale" advantages in regulation.

Also, the insolvency argument points in the direction of regulation. Pollution can be caused by individuals or firms with assets which are generally lower than the damages they can cause by the pollution. In this respect it should not be forgotten that even a small firm can cause harm to a large number of individuals or to entire ecosystems. The amount of damages caused by this emission can of course largely exceed his individual assets. Moreover, most firms have been incorporated as a legal entity and therefore benefit from limited liability. Hence, the individual shareholders are not liable to the extent of their personal assets, but a creditor of the firm can only lay claim to part on all of the total assets purchased in the firm by the shareholders.

Also the chances of a liability suit being brought for damage caused by wrongful pollution is naturally very low. The damage is often spread over a large number of people, who will have difficulties to organise themselves to bring a law suit. In addition, the damage could only

[6] These are often referred to as "public interest" criteria for regulation to contrast them with "private interest" explanations for regulation, as advanced by public choice scholars.

become apparent some years after the emission took place. This will bring proof of causation and latency problems, which will only make it difficult for a lawsuit to be brought against the polluter.

For these reasons it is clear that some form of government regulation of environmental pollution is necessary. To reformulate : this shows that liability rules alone can not suffice to prevent environmental harm, but there might be other, publicly imposed, instruments than the command and control type regulation which can be used to reach this goal. Taxes are obviously such an alternative. But also these are publicly imposed and can hence be considered as 'regulation' Another question, which will be discussed below, is whether this necessarily implies that environmental law should solely depend upon regulation or whether regulation can still fulfill a supplementary role.

2.4 Safety regulation in practice

When Shavell's criteria for safety regulation are applied to the environmental risk, one can easily note that a strong argument can be made that the efficient care to be taken to avoid environmental damage should also be fixed *ex ante* by regulation.

In many cases this regulation consists of licences or permits in which an administrative authority fixes an emission standard which must be followed by the potential polluter. These licences play a crucial role in environmental policy in most countries. An improvement of environmental quality will mostly be effected by imposing more stringent emission standards in administrative licences. Hence, the general requirement that emissions are controlled through licences and that the quality and quantity of the emissions are regulated by the conditions in this licence, is a cornerstone of environmental law. Since these licences are administrative acts, in most legal systems environmental law is considered to be a part of administrative law. Criminal law usually only comes into the picture to sanction a violation of administrative regulations or emission standards in the licences.

Although environmental pollution is in the first place controlled through these administrative licences, in individual cases there can still be damage to the environment. Then again liability under tort law comes into the picture and the question is raised of the influence of regulation on the liability system and *vice versa*[7]. These complimentarities between tort law and regulation shall be discussed below.

Although it is difficult to examine whether the environmental regulation is generally also effective in reducing environmental harm, some studies have attempted to examine the effectiveness of safety regulation in controlling environmental harm. These studies do not

[7] Complementarities between tort law and regulation have been addressed by Rose-Ackerman, 1991, 1992a and 1992b.

address the specific quality of every environmental law, but examine whether regulation has generally been more important in reducing environmental harm than liability rules. Dewees demonstrated that in North-America the quality of the environment has improved substantially as a result of regulatory efforts, not so much in response to legal action in tort (see Dewees, 1992).

This empirical evidence of the success of regulation, compared to tort law, has been stressed in the recent book of Dewees/Duff/Trebilcock (see Dewees, Duff and Trebilcock, 1996). They hold that the large regulatory effort to improve the environment has met with considerable success when measured by the reduction of emissions, but that it is more difficult to argue that the environmental regulations of the 1970's in U.S. equally had a considerable influence on the ambient environmental quality. Moreover, they also stress that while environmental regulation is a determining factor in pollutant emissions and ambient concentrations, other non-regulatory factors such as economic growth and even the weather also influence environmental quality (see Dewees, Duff and Trebilcock, 1996).

3 Liability and regulation combined

3.1 Necessity of the combination

I just stressed that according to Shavell's criteria there is a strong argument to control the environmental risk through *ex ante* regulation (or taxes). However, in individual cases there can still be damage to the environment. Then again liability under tort comes into the picture and the question has been addressed in the literature how regulation influences the liability system and vice versa. These complementarities between tort law and regulation have more particularly been addressed by Rose-Ackerman (see Rose-Ackerman, 1992b and 1996), Faure/Ruegg (see Faure and Ruegg, 1994) and Kolstad/Ulen/Johnson (see Kolstad, Ulen and Johnson, 1990). Rose-Ackerman also compared US and European experiences in using regulation versus tort law in environmental policy (see Rose-Ackerman, 1992a and 1995). The first point which is often stressed, is that the fact that there are many arguments in favour of *ex ante* regulation of the environment, does not mean that the tort system should not be used any longer for its deterring and compensating functions. One reason to still rely on the tort system is that the effectiveness of (environmental) regulation is dependent upon enforcement, which may be weak. In addition the influence of lobby groups on regulation, to which public choice theory has rightly pointed, can to some extent be overcome by combining safety regulation and liability rules. Moreover, safety regulation, e.g. emission standards in

licences, can be outdated fast and often lacks flexibility, which equally merits a combination with tort rules.

Hence, from the above it follows that although there is a strong case for safety regulation to control the environmental risk, tort rules will still play an important role as well. Hence, the question arises what the influence is of regulation on the liability system and vice versa. How do these two systems mutually influence each other ?

3.2 Violation of regulation and liability

The first question to be answered in that respect is whether a violation of a regulatory standard should automatically be considered a fault under tort law and thus lead to liability of the licensee.

Assuming that the licence sets the regulatory standard at the efficient care level a violation of the regulatory standard should indeed lead to liability to give the licensee an incentive to spend on care. However, Shavell argues that the costs of following the regulatory standard are not the same for all injurers. Following the standard might be inefficient for some injurers. The injurers for whom following the regulatory standard would only be possible at high costs should not be held to follow this standard since it would create inefficiencies (see Shavell, 1994, 365-366; Faure and Van den Bergh, 1987, 109-100). The question is whether this means that these injurers should not be held liable if they violate the regulatory standard.

This problem can be compared with the *bonus pater familias* standard used in tort law. Although a detailed individualisation of standards of efficient care would be the optimal solution in a first best world this is often impossible given the costs of an individualised standard setting. Therefore, the legal system sets the required level of care at an average level, the so-called *bonus pater familias* standard. The same can be said for regulation. If various groups can be identified at low costs a separate standard for a certain group is efficient as long as the gains from selecting a further group outweigh the further administrative costs. In most cases, however, the regulator will not have the possibility of identifying atypical parties that might be able to avoid a loss at lower costs, for instance because they pose lower risks than normal. Therefore, a single regulatory standard will be used (see Posner, 1998, 183-184; Shavell, 1987, 74).

Although one could, therefore, argue that a failure to satisfy the regulatory requirement should not necessarily result in a finding of negligence, so as to avoid some parties who pose lower risks taking wasteful precautions (see Shavell, 1984, 365-366), most legal systems generally consider a breach of a regulatory duty a fault. One of the reasons for introducing safety regulation to prevent environmental damage is, as was mentioned above, that the regulator will usually pos-

sess better information to evaluate the efficient standard of care than the parties involved. Hence, the regulation passes on information to the parties on the efficient standard of care. The regulation also gives information to the judge who has to evaluate the behaviour of the injurer *ex post* in a liability case. The judge might lack the information necessary to find out whether in a particular case an injurer should not be held to follow the regulatory standard, for example because he posed a lower risk than usual. Therefore, particularly in environmental cases a judge will accept a finding of negligence as soon as a regulatory standard has been breached[8]. Thus, the statutory standards can be applied to define negligence (see Rose-Ackerman, 1992a, 127).

3.3 Compliance with regulation and liability

Whereas according to tort law in many legal systems a breach of a regulatory standard results automatically in a finding of negligence, the opposite is not true : following a regulatory standard does not exclude a finding of liability. In environmental law this is particularly important, since the conditions under which an emission of pollutants is allowed are mostly laid down in a permit. The industry often argues that as long as they follow the conditions of the licence, no finding of negligence in tort law is possible.

This point of view is, however, firmly rejected in most legal systems, for instance in both Belgium and in the Netherlands[9]. The basic idea is that the administrative authority, when granting a licence and setting permit conditions, cannot take into account the possible harm that the licensed activity might cause to all possible third parties. Their rights on compensation for damages may not be impaired simply because the operator of a plant followed the conditions of a licence. Legal doctrine and case law clearly state that keeping the permit conditions is just a minimum; in addition, the plant owner has to take all possible precautions as deemed necessary under tort law to avoid his licensed activity causing harm to third parties.

For instance in Dutch case law it is indeed generally accepted that following the conditions of a licence does not release a plant owner from potential liability [10]. An exception would only exist if the interests of the potential victims were clearly taken into account when the conditions

[8] Faure and Van den Bergh have also argued that an advantage of this system is that it gives victims incentives to prove that the regulatory standard has been breached. This makes the victim an enforcer of safety regulation. He can claim compensation under the negligence rule as soon as a causal relationship between the violation of the regulatory standard and his damage is established. See Faure and Van den Bergh, 1987, 110-111.

[9] For a comparative analysis of the question whether following a permit excludes criminal liability see, Faure and Oudijk, 1994, 86-91.

[10] See Supreme Court (Hoge Raad) 30 January 1914, *Nederlandse Jurisprudentie*, 1914, 497; Supreme Court 10 March 1972, *Nederlandse Jurisprudentie*, 1972, 278.

of the permit were set (see Rus-van der Veld, 1987, 111; Nieuwenhuis, 1991, 44-47). This point is made very clear in a famous case in the Dutch Supreme Court that dealt with pollution caused by the French salt mines in the Alsace region[11]. The Salt Mines argued that the emissions were within the limits set by their permit and, therefore, not illegal. The court, however, judged that the licence had not taken into account the potential harmful effects of the emissions for third parties and could, therefore, not release the salt mines from liability.

One can find a clear economic rationale for this rule. If compliance with a regulatory standard or licence would automatically result in a release from liability, the potential injurer would have no incentive to invest more in care than the regulation asks from him, even if additional care could still reduce the expected accident costs beneficially (see Shavell, 1984, 365; Faure and Van den Bergh, 1987, 110). A first reason to hold an injurer liable (if the other conditions for liability are met), although he has followed the regulatory standard, is that this standard is often merely a minimum. Exposure to liability will give the potential injurer incentives to take all efficient precautions, even if this requires more than just following the licence. A second reason is that exposure to liability might be a good remedy for the unavoidable capturing and public choice effects that play a role when permits are granted. If a permit would always release from liability, all a plant operator would have to do, is get a good permit with easy conditions from a friendly civil servant. That would then exclude any law suit for damages from a potential victim. Finally, tort law can also be seen as a 'stopgap' for situations not dealt with by the statute (see Rose-Ackerman, 1992a, 123). This makes clear that the exposure to liability notwithstanding the permit is an important guarantee that the plant operator will take efficient care.

Therefore, following the conditions of a license or – more generally – regulatory standards, should not have a justificative effect in tort. The opposite may only be true if it were clear that the administrative agency took into account all potential harm of all interested third parties when setting permit conditions. In such case a judge in an civil liability suit should not be "second guessing" efficient agency decisions. It is, however, rare that agencies will be able to take *ex ante* all these interests and possible damages into account when setting permit conditions. Hence, as a general rule, following licenses or regulatory standards should not free from liability; the opposite would be the exception. This is the case both under a negligence as well as under a strict liability rule. Indeed, holding an injurer liable, notwithstanding he followed regulatory standards will play an important role under a strict liability rule, since this will lead the injurer to take efficient care and adopt an efficient activity level, i.e. to take all efficient measures to reduce the potential accident

[11] Supreme Court 23 September 1988, *Rechtspraak van de Week*, 1988, 150 and see Faure, 1991, 128-129.

costs, although this might require more to be done than the regulation requires. Under a negligence rule this case law is also significant if the efficient care standard (which is assumed to be equal to the due care standard required by the legal system) is higher than the regulatory standard.

4 Environmental liability

Let us finally examine how exactly environmental liability could be used as a tool to prevent environmental harm.

4.1 Negligence versus strict liability : economic principles

One possible liability rule which will give the polluter an incentive to spend on care to reach the optimal standard is the negligence rule. This follows from the general literature on the economics of accident law (see Shavell, 1987, 8 and Calabresi, 1975, 658). Assuming that under a negligence rule the potential polluter will only have to pay compensation if he spends less on care than the legal system wants him to (due care) the firm will have an incentive to spend on care, since it is a way to avoid liability which will maximise his utility. Provided that the legal system defines the due care level as the optimal standard a negligence rule will therefore give the polluter incentives to follow the optimal standard. Also a strict liability rule will lead to optimal incentives for care taking for the polluter, since taking efficient care will minimise the expected accident costs which the potential polluter has to bear under a strict liability system (see Polinsky, 1983, 39; Shavell, 1980, 11 and 1987, 8). Therefore, the literature generally accepts that both a negligence rule and a strict liability rule will provide a potential polluter with incentives to take the efficient care level. However, this is only valid in a unilateral accident setting, i.e. an accident whereby only the injurer can influence the accident risk. If victims were also to be given incentives for accident reduction a contributory negligence defence should be added to the strict liability rule. Under negligence victims will always have an incentive to take efficient care as well since they will in principle not be compensated by the injurer who, under a negligence rule, will take efficient care to avoid liability.

However, the accident risk is not only influenced by the level of care, but also by the number of times that the parties are involved in the risky activity, i.e. the activity level. Hence, an optimal liability rule should also give the parties in a potential accident setting incentives to adopt an optimal activity level. A negligence rule will not give optimal incentives to the injurer to adopt an optimal activity level since the activity level is not incorporated in the due care standard which

the court applies (see Adams, 1989; Diamond, 1974; Shavell, 1980). Hence, under a negligence rule the injurer only has an incentive to take efficient care (to escape liability) but not to adopt an efficient activity level. Under a strict liability rule, on the contrary, an injurer has an incentive to adopt an efficient activity level since this is also a way to minimise the total expected accident costs which he has to bear. Moreover, under negligence the injurer will only take the due care, the legal system requires from him, but he has no incentive to take other precautionary measures which could reduce the risk of environmental damage. Strict liability has the advantage, that it gives incentives to the injurer to take all efficient precautionary measures to reduce the risk, also those which could not be incorporated into the due care level under negligence. In a unilateral accident model (whereby only the behaviour of the injurer influences the accident risk) strict liability is therefore the efficient liability rule since it leads both to efficient care and to an optimal activity level.

4.2 Legal justifications for strict liability

The reason that is often advanced in legal literature in favour of strict (environmental) liability is that strict liability will help the victim in obtaining compensation since he is released from the heavy burden of proving fault under the negligence rule. However, from a deterrency point of view victim compensation is not as such a goal of accident law. The duty of the injurer to compensate his victim is only an instrument to reach deterrence efficiency. Moreover, the victim compensation argument to introduce strict liability for environmental pollution is not that convincing in all cases. Indeed, many legal systems qualify every violation of a statutory or regulatory norm as a civil fault[12]. Most industries are subjected to extensive safety regulation. Hence, in these systems the victim only has to prove the violation of one of these regulations to establish a fault[13]. If, in addition, the victim can prove a causal relationship with the loss suffered, he will be able to claim compensation. In many accident cases this burden of proof will therefore not be as heavy as has been argued. It is, therefore, at least questionable whether a strict liability rule substantially improves the situation of the victim in comparison with an already existing broadly interpreted civil fault regime. It should also not be overlooked that under the general fault regime of tort law no limitations apply and the victim is entitled to full compensation. In many of the environmental cases where strict liability was first introduced, more particularly in the international conventions concerning nuclear accidents and oil pollu-

[12] See for Dutch case law e.g. the so called Jumbo II case of 1 October 1993, *Nederlandse Jurisprudentie*, 1995, 182.

[13] The economic rationale for this rule was discussed above.

tion[14], financial caps and other limitations on the victim's rights were introduced. The alleged compensating benefit of the strict liability in those cases is therefore doubtful.

4.3 Strict liability for environmental damage ?

Although the classic victim compensation argument may as such not justify the introduction of strict liability for environmental pollution, there are on the other hand economic reasons based on deterrence efficiency for introducing a strict liability rule. Environmental pollution can in most cases certainly be considered a unilateral accident, i.e. an accident whereby only the injurer can influence the accident risk. In this case we noted that the economic model predicts that the advantage of the strict liability rule is that it will give the injurer an incentive both to adopt an optimal activity level and to take efficient care. Since the victim cannot influence the accident risk, strict liability seems to be the first best solution to give the potential polluter optimal incentives for accident reduction in those cases (see Faure, 1995).

In sum, if we apply the criteria of Shavell determining the choice between negligence and strict liability to the environmental case, there seem to be strong arguments in favour of an introduction of strict liability. In many cases environmental pollution will be truly unilateral in the sense that only the injurer's activity can influence the accident risk, which constitutes a strong case for strict liability[15]. In other cases the victim will certainly be able to exercise an influence on the risk as well. One can more specifically think about situations where the victim has the possibility to mitigate damages after the accident occurred. However, in those cases it is not the victim's activity level, but his level of care which influences the accident risk. This can be controlled by adding a contributory or comparative negligence defence to the strict liability rule.

4.4 A few refinements

Many scholars argued that there is indeed a strong case in favour of strict liability for environmental damage : this will give the potential polluter optimal incentives for accident reduction and hence, for optimal internalisation (see Endres and Staiger, 1996). There is, however, another important aspect of the difference between negligence and strict

[14] See with respect to nuclear accidents OECD, 1994; Faure and Skogh, 1992, 499-513; Deprimoz, 1995, 1-24 and with respect to civil liability for marine oil pollution Faure and Heine, 1991, 39-54 and for recent evolutions Brans, 1994, 61-67 and 85-91.

[15] In some cases it will be the victim's activity that caused the harm, e.g. if the victim knowingly came to the nuisance. This may then lead to a denial of a claim on compensation. See in that respect the discussion on the coming to the nuisance doctrine, by Wittman, 1980, 557-568.

liability which should be mentioned. This concerns the fact that the application of negligence requires high information costs from the judge, who will have to set the due care standard. The information necessary to weigh costs and benefits and fix the optimal care may not be readily available for the judge. Strict liability shifts all costs to the injurer, who will then have to define the optimal care level. If one therefore assumes that, as may be the case with environmental harm, the information on optimal precaution is better available with industry than with the judges, this constitutes an argument for strict liability. This information advantage may therefore constitute an additional argument in favour of strict liability for environmental harm. One should, however, remember that this finding only holds in all the models, such as the one which is e.g. developed by Shavell, which start from an assumption of risk neutrality. If risk aversion is introduced and the potential injurer is risk averse, Endres/Schwarze correctly argue that strict liability is only efficient if in some way risk can be removed from the risk averse injurer, e.g. through insurance (see Endres and Schwarze, 1991).

There are other reasons why the seemingly advantage of strict liability should be somewhat balanced. First of all, it was assumed until now that the injurer has money at stake to pay compensation to the victim. If, however, the amount of the damage exceeds the injurer's wealth, a problem of underdeterrence will arise. Under strict liability the injurer will consider the accident as one which is equal to his total wealth and will therefore only take the care necessary to avoid an accident with a magnitude equal to his total wealth. If that wealth is lower than the magnitude of an accident he will take less than the optimal care and therefore a problem of underdeterrence arises under strict liability. Insolvency is less of a problem under negligence since under that rule the injurer will still have an incentive to take the care required by the legal system as long as the costs of taking care are less than his individual wealth. Taking due care remains indeed a way for the injurer to avoid to have to pay compensation to the victim. If there would thus be a potential accident setting whereby the magnitude of the loss may be higher than the injurer's wealth (which can often be the case in environmental liability) this constitutes an argument in favor of negligence rather than strict liability.

4.5 White paper on environmental liability

As we already indicated, in many international conventions, a strict liability regime is introduced for environmental harm. This is equally the case in many legal systems. Recently the European Commission introduced a white paper on environmental liability (see COM (2000) 66 final of 9 Februari 2000). In that white paper the Commission took a rather balanced approach towards the choice between negligence and strict liability. The Commission opts for a strict liability rule for all

harm which originates from hazardous activities. For all the damage originating from other activities the Commission proposes a negligence rule.

This original approach of the European Commission is very much in line with the predictions of the economic model as presented above. Hazardous activities can often be considered unilateral and in those situations it is important to control the injurer's activity through a strict liability rule. The same is, however, not true in case of non-hazardous activities which may cause environmental harm. The dividing line (non-hazardous/hazardous) chosen by the European Commission, seems therefore to follow economic logic.

5 Concluding remarks

In this paper I tried, without using formal models or equations, to prove the importance of traditional concepts of law and economics for environmental lawyers and policy makers. By focusing on various aspects of environmental law (more specifically liability, regulation and the combination of the two) the practical use of the economic analysis of law was demonstrated. Law and economics may thus contribute towards a better understanding of environmental law, e.g. by providing an understanding of why environmental law is so often subjected to regulation. In addition, the paper showed that traditional law and economics can also contribute to environmental economics. Law and economics e.g. points at the deterrent effect of liability rules, a point which is often neglected in traditional writings on environmental economics. Moreover, law and economics stresses the importance of a combined use of various instruments to control environmental pollution by pointing at strength and weaknesses of each legal instrument, taken separately.

A rather reassuring conclusion for environmental lawyers is that current environmental law in many Western European legal systems seems to correspond to a large extent to economic logic. However, the fact that environmental law is e.g. relying heavily on regulation may correspond with Shavell's criteria for safety regulation, it does on the other hand not imply that every specific type of environmental regulation is always efficient. Indeed, there may be various reasons why regulatory outcomes may fail to be effective. One reason may be poor information by the regulator; another one concerns the risk that the regulator is captured by special interests. For these and many other reasons, some of the traditional (command and control) regulation in environmental law has often been considered inefficient by economists.

Moreover, the analysis presented in this paper could only provide a basic introduction to the economics of environmental law. There are,

however, various caveats which have to be considered :

This paper mainly adressed the question what instruments should be used to prevent environmental harm. It did, however, not concern the question to what level environmental pollution should be internalized. Obviously environmental harm should, from an economic point of view, not be prevented at all costs, but should imply some weighing of costs and benefits.

Second, I only focused on two specific, be it important, policy instruments to control environmental harm, regulation and liability. Many other instruments which become increasingly important today, such as taxes and more particularly market oriented instruments such as marketable permits have not been discussed. Future environmental law and economics research will focus on the possible combined use of various of those instruments.

Third, in this paper it was assumed that external force (e.g. the threat of being held liable) is necessary to give a potential polluter incentives to prevent environmental harm. However, there may be a number of reasons why companies would engage voluntarily in investments in environmentally friendly technology. In some cases a change to cleaner production technologies may lead both to reduction of environmental harm and to a better economic performance of the firm concerned. In addition, a firm may find it advantagous to present itself as "green" as a marketing tool. Thus it could e.g. submit itself voluntarily to eco-audits. Thus there may be many (economic) reasons for the voluntary implementation of technologies which lead to environmental improvements, without the necessity to have a legal rule coercing towards environmental improvement.

Finally it is important to remember that economics, how undoubtedly useful it can be, provides only "one view of the cathedral" (compare Calabresi and Melamed, 1972) by analysing legal rules on the basis of the efficiency criterion. Environmental lawyers and policy makers may, however, sometimes have other goals they want to achieve (at least on paper) than economic efficiency. But even if policy makers would e.g. define the goal of environmental policy as being something like "environmental justice", than economic analysis remains a useful tool. It can indeed explain to policy makers how to obtain the maximum "environmental justice" per dollar spend[16].

[16] Compare Easterbrook, 1983, 289-332 who argues the same with respect to those who would reject a deterrence-based approach to criminal law.

References

Adams, M., 1989, "New Activities and the Efficient Liability Rules", in M. Faure. and R. Van den Bergh (eds), *Essays in Law and Economics. Corporations, Accident Prevention and Compensation for Losses*, Maklu.

Bocken, H., 1987 and 1988, "Alternatives to Liability and liability insurance for the compensation of pollution damages", *Tijdschrift voor Milieuaansprakelijkheid*, 4-87, pp. 83-87 and *Tijdschrift voor Milieuaansprakelijkheid*, 1-88, pp. 3-10.

Brans, E. H. P., 1994, "Liability for Ecological Damage under the 1992 Protocols to the Civil Liability Convention and the Fund Convention and the Oil Pollution Act of 1990", *Tijdschrift voor Milieuaansprakelijkheid*, pp. 61-67 and pp. 85-91.

Calabresi, G. and Melamed, D., 1972, "Property rules, liability rules and inalienability : one view of the cathedral", *Harvard Law Review*, 85, pp. 1089-1128.

Calabresi, G., 1975, "Optimal Deterrence and Accidents", *Yale Law Journal*, 84, pp. 656-671..

Coase, R.H., 1960, "The Problem of Social Cost", *Journal of Law and Economics*, 1 (1), pp. 1-44.

Deprimoz, J., 1995, "Regime Juridique des Assurances contre les Risques Nucléaires", *Juris. Classeur*, 555, pp. 1-24.

Dewees, D, Duff, D. and M. Trebilcock, 1996, *Exploring the Domain of Accident Law. Taking the Facts Seriously*, New York, Oxford, Oxford University Press.

Dewees, D., 1992, "The Comparative Efficacy of Tort Law and Regulation for Environmental Protection", *Geneva Papers on Risk and Insurance*, pp. 446-467.

Dewees, D., 1992, "Tort Law and the Deterrence of Environmental Pollution" in Tietenberg, T.H., (ed.), *Innovation in Environmental Policy, Economic and Legal Aspects of Recent Developments in Environmental Enforcement of Liability*, Brookfield, Elgon, pp. 139-164.

Diamond, P., 1974, "Single Activity Accidents", *Journal of Legal Studies*, pp. 107-164.

Easterbrook, F., 1983, "Criminal procedure as a market system", *Journal of Legal Studies*, pp. 289-332.

Endres, A. and B. Staiger, 1996. "Ökonomische Aspekte des Umwelthaftungsrecht", in M. Ahrens and J. Simon, (eds.), *Umwelthafung, Risikosteuerung und Versicherung*, Berlin, Ehrich Schmidt Verlag, pp. 79-93.

Endres, A. and R. Schwarze, 1991, "Allokationswirkungen einer Umwelthaftpflicht-versicherung", *Zeitschrift für Umweltpolitik und Umweltrecht*, pp. 1-25.

Faure, M. and G. Skogh, 1992, "Compensation for Damages Caused by Nuclear Accidents : A Convention as Insurance", *The Geneva Papers on Risk and Insurance*, pp. 499-513.

Faure, M. and G. Heine, "The Insurance of Fines : The Case of Oil Pollution", *The Geneva Papers on Risk and Insurance*, 1991, pp. 39-54.

Faure, M. and J.-C. Oudijk, 1994, "Die strafgerichtliche Überprüfung von Verwaltungsakten im Umweltrecht. Ein rechtsvergleichender Überblick der Systeme in Deutschland, den Niederlanden und Belgien", *Juristenzeitung*, pp. 86-91.

Faure, M. and M. Ruegg, 1994, "Standard Setting through General Principles of Environmental Law", in M. Faure, J. Vervaele and A. Weale. (eds.), *Environmental Standards in the European Union in an Interdisciplinary Framework*, Antwerp, Maklu, pp. 39-60.

Faure, M., 1991. "De gevolgen van de 'administratieve afhankelijkheid' van het milieustrafrecht : een inventarisatie van knelpunten" in M. Faure, J. Oudijk and D. Schaffmeister (eds.), *Zorgen van Heden*, Arnhem, Gouda-Quint, pp. 128-129.

Faure, M., 1995, "Economic Models of Compensation for Damage Caused by Nuclear Accidents : Some Lessons for the Revision of the Paris and Vienna Conventions", *European Journal of Law and Economics*, pp. 21-43.

Faure, M., 1996, "Economic aspects of environmental liability : an introduction", *European Review of Private Law*, pp. 85-109.

Gunningham, N. and P. Grabosky, 1998, *Smart Regulation. Designing Environmental Policy*, Oxford, Clarendon Press.

Kolstad, Ch.D., T. S. Ulen and G. V. Johnson, 1990, "*Ex Post* Liability for Harm vs. *Ex Ante* Safety Regulation : Substitutes or Compliments ?", *American Economic Review*, 80, pp. 888-901.

Landes, W. and R. Posner, 1984, "Tort law as a regulatory regime for catastrophic personal injuries", *Journal of Legal Studies*, pp. 417-434.

MacKay, E., 1982, *Economics of information and the law*, Boston, Kluwer.

Nieuwenhuis, J.H., 1991, "Blinddoek en balans in het milieurecht. Drie manieren om belangen af te wegen" in A. M. Hol and M. A. Loth (eds.), *Dilemma's van aansprakelijkheid*, pp. 44-47.

OECD, 1994, *Liability and Compensation for Nuclear Damage, An International Overview*, Paris, OECD.

Polinsky, A.M., 1983, *Introduction to Law and Economics*, Boston, Little, Brown & Co.

Posner, R., 1998, *Economic Analysis of Law*, 5th Edition, New York, Aspen.

Rose-Ackerman, S., 1991, "Regulation and the Law of Torts", *American Economic Review, Papers and Proceedings*, May, pp. 54-58.

Rose-Ackerman, S., 1992 a, *Re-thinking the Progressive Agenda, the Reform of the American Regulatory State*, New York, The Free Press.

Rose-Ackerman, S., 1992 b, "Environmental Liability Law", in T. H. Tietenberg (ed.), *Innovation in Environmental Policy, Economic and Legal Aspects of Recent Developments in Environmental Enforcement and Liability*, Brookfield, Edward Elgar, pp. 223-243.

Rose-Ackerman, S., 1995, "Public Law versus Private Law in Environmental Regulation : European Union Proposals in the Light of United States Experience", *Review of European Community and International Environmental Law, RECIEL*, 4, pp. 312-32.

Rose-Ackerman, S., 1995, *Controlling Environmental Policy : the Limits of Public Law in Germany and the United States*, New Haven and London, Yale University Press.

Rose-Ackerman, S., 1996, "Public Law versus Private Law in Environmental Regulation : European Union Proposals in the Light of United States and German Experiences", in Eide, E. and Van den Bergh, R. (eds.), *Law and Economics of the Environment*, Oslo, Juridisk Forlag, pp. 13-39.

Rus-van der Veld, M., 1987, "Onrechtmatige daad en vergunning", in N.S.J. Koeman, W. J. Ouwerkerk and J. M. Van Dunné (Eds.), *Civielrechtelijke aansprakelijkheid voor milieuschade*, Vereniging voor Milieurecht, 4, Tjeenk Willink, Zwolle, pp. 105-113.

Schwartz, A. and L. Wilde, 1979, "Intervening in markets on the basis of imperfect information : a legal and economic analysis", *University of Pennsylvania Law Review*, pp. 630-682.

Shavell, S., 1980, "Strict Liability versus Negligence", *Journal of Legal Studies*, pp. 1-25.

Shavell, S., 1984 a, "A Model of the Optimal Use of Liability and Safety Regulation", *Rand Journal of Economics*, pp. 271-280.

Shavell, S., 1984 b, "Liability for Harm versus Regulation of Safety", *Journal of Legal Studies*, pp. 357-374.

Shavell, S., 1985, "Criminal law and the optimal use of non-monetary sanctions as a deterrent", *Columbia Law Review*, pp. 1232-1262.

Shavell, S., 1986, "The judgement proof problem", *International Review of Law and Economics*, pp. 43-58.

Shavell, S., 1987. *Economic Analysis of Accident Law*, Cambridge, Harvard University Press.

Stigler, G., 1961, "The economics of information", *Journal of Political Economy*.

Van den Bergh, R., 1988, "Le droit civil façe à l'analyse économique du droit", *Revue Internationale de Droit Economique*.

Wittman, D., 1980, "First come, first served : an economic analysis of coming to nuisance"', *Journal of Legal Studies*, pp. 557-568.

Résumé

Cet article montre comment différents d'instruments juridiques peuvent être utilisée pour traiter les problèmes de pollution environnementale. On démontre pourquoi les règles de responsabilité ne peuvent suffire pour contrôler les risques environnementaux. Le recours à la réglementation ne peut non plus suffire. Il faut donc mélanger des règles de responsabilité et la réglementation. Comment ? Telle est la question que résoud cet article. Une autre question est également abordée : quel est le type de responsabillité optimal afin d'internaliser les risques environnementaux.

Abstract

The purpose of this paper is to show how a variety of traditional legal instruments can be used to remedy environmental pollution. Using the traditional public interest criteria for regulation, it is explained why liability rules alone can not suffice to control the risks posed by environmental harm. However, the conclusion that in environmental law some regulatory invention will be necessary does not exclude the role of environmental liability. Hence the question arises how liability rules and regulation can be used jointly to remedy environmental pollution and how they mutually influence each other. Finally, the question will have to be asked which of these liability rules is optimal to internalize environmental risks.

Mots clé

Environnement, responsabilité civile, réglementation, responsabilité sans faute

Key words

Liability, regulation, strict liability, environmental liability

Classification JEL : K00, K13, K32

Economic Analysis of Environmental Laws and Regulations

Jashaswi Ghosh[1]

ABSTRACT. In the past century there has been a great deal of deliberation and research regarding the state of the environmental degradation which has been a steady and ever increasing phenomenon enhanced by the industrial revolution which has spread across the length and breadth of this planet. Majorly irresponsible human actions have been held responsible for the destruction of the environment. Thus, the countries are now trying to improve this situation by entering into multi-lateral agreements and so on, which are aimed at reducing pollution. However, this paper has tried to analyse how truly efficient these laws and regulations are and tries to point out their inefficiencies and deficiencies. This paper by means of using economic theories and tools shows how the international laws and regulations are economically inefficient as well as inefficient in terms of regulating the degradation that is caused to the environment. Thus basically, the aim of the paper is to point out the flaws and laxes in the international regulations and show the resulting economic losses by means of tax fraud, carbon credit recycling and so on and finally suggesting certain changes in the structure of the laws which will make them more efficient and sound.

1 Student, 2nd Year, B.A., LL.B (Hons.), Gujarat National Law University, Gandhinagar; The author may be reached at kasturyghosh@gmail.com.

Published in Articles section of www.manupatra.com

CONTENTS

INTRODUCTION 77
INCOME EFFECT AND DEMAND FOR ENVIRONMENTAL QUALITY 78
REGULATORY APPROACHES AND ITS DEFFICIENCY 81
INTERNATIONAL TRADE 82
TAX FRAUD IN INTERNATIONAL TRADE 85
CONCLUSION AND SUGGESTIONS 87

Published in Articles section of www.manupatra.com

INTRODUCTION

In the past century, there has been a great deal of deliberation and research regarding the state of environmental degradation which has been a steady and ever increasing phenomenon enhanced by the industrial revolution which has spread wide across the length and breadth of this planet. Scientific evidence has always pointed to the irresponsible human activities and industrial pollution and waste as a primary reason of this degradation. The field of modern international environmental law has its origins in a dispute between the United States and Canada over air pollution damages in Washington State from an ore smelter in Trail, British Columbia.[2] The resolution of this dispute established the norm of customary international law that it is the State's duty to avoid letting its activities produce harm in other States. Failing that duty, a State is liable to compensate for environmental damages. International environmental law in the 1950s and 1960s proceeded almost exclusively in the form of such customary laws. In the 1970s, treaties began to codify the customary norms of international environmental law. The 1980s and 1990s witnessed a proliferation in the magnitude, complexity, and scope of international environmental law treaties. No longer content to merely codify existing norms, these new treaties imposed new duties and standards on States and other actors.[3] In response to this scientific evidence, the global community agreed in 1992 to an international treaty, the United Nations Framework Convention on Climate Change (UNFCCC). The treaty requires countries to cooperatively consider what they could do to limit average global temperature increases and the resulting climate change and to cope with whatever impacts were, by then, inevitable. As at June 2013, the treaty has been ratified by 195 parties.[4]

A system of carbon credits was established. A very clear record of this carbon credit trade transaction is maintained by the United Nations Climate Change Secretariat, in Bonn, Germany for not only keeping records but also for checking whether these transactions are following the procedures prescribed by the Kyoto Protocol. In spite of the regulatory procedures which are followed in the "carbon market" to maintain the legality of the trade, several counterfeiting methods are being used by potential traders in this market to increase their profits exponentially and thus enhancing the environmental degradation procedure. INTERPOL, through its Environmental Crime Programme and the Economic and Financial Crimes sub-Directorate, recognized that emerging carbon markets, like any market, are at risk of exploitation through criminal means and therefore require proper monitoring and enforcement to ensure environmental and financial integrity.[5] Thus, we see that linkages between the international trading regime and efforts to address

2 3 Trail Smelter Arbitration (*United States v. Can.*), 3 R.I.A.A. 1905 (1941), reprinted in 35 Am. J. Int'l L. 684 (1941).
3 Huang, Peter H. "International Environmental Law and Emotional Rational Choice." *The Journal of Legal Studies*: S237-258.
4 Status of Ratification of the Convention, UNFCCC <http://unfccc.int/essential_background/convention/status_of_ratification/items/2631.php>, Last accessed on 10th September 2014.
5 INTERPOL GUIDE TO CARBON TRADING CRIME (2013).

Published in Articles section of www.manupatra.com

climate change require governments and policy makers in each area to take account of the other. Trade rules may both constrain and facilitate action to address climate change, while actions taken to address climate change will in some cases have important implications for trading relationships.

Thus to address several problems which are arising in the carbon market we have to begin by knowing the kind of problems which are existing and then try to find out methods of solving these problems. Moreover, in this paper we are also going to deal with the flaws in the environmental rules and regulations and thus try to point out the basic deficiencies which are present in them and try to find ways to amend these flaws.

INCOME EFFECT AND DEMAND FOR ENVIRONMENTAL QUALITY

The answer to the question as to why developing countries tend to have more lax environmental regulations than rich countries, comes from the very basic effect that income seems to have on the demand for environmental regulations and quality. As people become wealthier, they seem to demand more and more environmental protection. Conversely, at any point in time, the poor seem to demand less protection than the rich.[6]

We know that environmental quality is not a homogeneous good, for example we have different requirements of these goods like quality of water, ambient air quality, and global climate and so on. However, the major question that is to be scrutinised and answered is, how does the demand for these goods increase with an increase in the income. As the world becomes wealthier, the extent to which environmental protection is to be provided becomes more or less dependent on the income elasticity of demand for the environmental quality. Thus we can clearly explain the differences in pollution levels in the various countries of the world by means of the differences which are existing in their income.

This is because of the simple reason that individual demand for the environmental quality will depend upon their own income. If a person does not have a high income, he will demand for goods which are cheaper so that he can afford them. An aggregate of these individual demands will represent the demand from the particular society. The government will act upon such demands and if the government is democratically elected, such demand will be the basis of the government's regulations. Thus, we can see how income is implicitly interlinked with the regulations which a particular government applies on its industries. Thus, it is logically assumed from the above information that the countries which are developing or are poor are not interested in their environmental quality and pollution, because their economic condition is not such that they can sustain a society which has stringent economic regulations which will not be favourable for the industries which are providing the economic opportunities to their society.

6 Kolstad, Charles D. *Environmental Economics*. New York: Oxford University Press, 2000. 246.

Published in Articles section of www.manupatra.com

Thus, their primary aim is to sustain and improve their economic condition rather than improve the quality of the surrounding environment.

The Kyoto protocol also recognises the plight of these developing countries. Simon Kuznets developed curves showing how income inequality changed as per capita income of the country increases. He found that income inequality initially increased and then decreased. Because of the similarity, curves showing how environmental quality or pollution change with changes in income in a country are known as **environmental Kuznets curves.**[7] This helps to determine policy questions which deal with improvement of environmental quality in poor countries.[8]

The competitive aspect in a market economy is of vital importance. Just as firms compete against each other by providing better goods, cheaper prices and so on, the Governments also compete against each other by providing incentives to the industries in order to attract these industries to establish their factories in the jurisdiction of the Government. They do this in a number of ways such as, providing tax reductions, cheap land for building the factories and so on. The governments have their own incentives in providing such amenities, like job opportunities for its citizens, economic boost and so on. However, one of the most important problems arising out of this is that of environmental degradation. Though this process is economically efficient at the first glance, in the long run, it has severe detrimental effects on the environment. Thus at the international level regulations vary substantially from one country to another. A simple reason for this is that environmental quality is a normal good, the wealthier people are, the more environmental quality they demand. Thus we would expect the citizens of a rich country to demand more environmental quality than citizens of a poor country. However, one negative impact of such process is the creation of "pollution Havens', which are countries that are full of polluting factories and industries because such countries have extremely relaxed environmental regulations and so on.

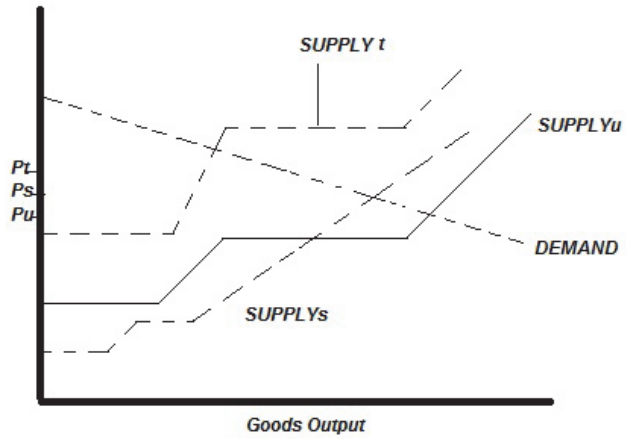
According to Coase Theorem[9] it makes no difference whether the polluter must compensate the victim of the pollution or the victim must pay the polluter not to pollute. There is an analogous issue in the context of Pigovian fees, which can be defined as a fee paid by the polluter per unit of pollution exactly equal to the aggregate marginal damage caused by the pollution when evaluated at the efficient level of pollution. The fee is generally paid to the government.[10] Thus a major question which arises out of this context is, whether the outcome of subsidizing firms is the same as the outcome obtained by imposing fees or taxes? This is an important question since subsidies are usually much more politically popular than taxes or fees[11]

7 *Ibid.* at pg. 248.
8 *Ibid.* At pg. 249.
9 *Ibid.* at pg. 124.
10 *Ibid.*
11 *Ibid.*

Published in Articles section of www.manupatra.com

We are determining what the market price of the good might be under these three regimes that is, taxes, subsidies and unregulated. The figure (figure A) illustrated below traces out the short-run supply functions for an industry for three cases, unregulated, Pigovian taxes and subsidies. Also shown is a typical demand function for the good. Recall that firms will operate on the portion of their marginal cost curve that lies above the variable cost. We can see that with both the unregulated and subsidy cases, both types of firms are operating, yielding product prices of Pu and Ps, respectively. In the case of Pigovian tax, only the newer firms operate, yielding product price Pt. Thus, from the graph it can be interpreted and can be concluded that taxes and subsidies have different effects in the short run. A subsidy may allow firms to continue operating which would not have happened in the case of tax. The subsidy requires a lump-sum transfer, which has to be obtained from somewhere. More importantly, the subsidy involves the operation of firms that are really losing money. This is not efficient.[12] Thus, from the above analysis it can easily be concluded that though the theoretical concept of Coase theorem tells us that subsidies and taxes have the same impact upon efficiency, we see that in case of practical application, as illustrated above, it has a detrimental effect, not only on the economy but more importantly on the environment by allowing the loss making and polluting firms to continue.

Figure A[13]:



The aspect of subsidy is undesirable because it does not allow the market to communicate the true costs of the product being consumed to the consumer. Thus suppose there are two factories, one of which makes goods from recycled materials, while the other is the polluting industry. Now, subsidies would basically require giving money to the polluting industry to enable it to clean up. This, would

12 *Ibid.* at pg. 127.
13 *Ibid.*

Published in Articles section of www.manupatra.com

however create an interest or incentive in investing in those polluting industries rather than industries which are using the recycled materials. Moreover as stated above we would want the consumers to see the full costs associated with producing the product so that the consumers know, what to buy and how much of it to buy.

To summarize, we could clearly suggest that, if jurisdictions/countries offer foreign firms tax-free status, it would not be in the best interest of these jurisdictions to loosen environmental regulations to attract capital. If, on the other hand, conditions are such that the jurisdiction must tax capital, we might see the complementary use of weak environmental regulations to make up for the capital taxation. In reality, we might expect positive capital taxation to exist in countries that are not excessively worried about attracting capital. It seems logical that such countries would not feel strongly enough about attracting capital that they would resort to substantial weakening of environmental regulations. This is just speculation, however; what happens in reality is an empirical question.

REGULATORY APPROACHES AND ITS DEFFICIENCY

The regulatory approach seeks to reach a given quality target for the environmental system by regulating individual behaviour. The general kinds of permits used to control environmental pollution or to regulate such pollution, are the pollution permits that are issued by the states which want to use such a mechanism. Now, these pollution permits are issued till the specified level of pollution that is pre-decided as the maximum level of pollution is reached. After that no further pollution permits are issued. However, caps on the amount of pollution produced by the individual countries or within them the individual industries are also a way of reaching a specified level of pollution and thus under the ambit of a regulatory approach. Thus, in several countries this approach gets widely used in environmental policy. Thus, water and air quality management in the United States is based on a permit system. Air-quality policy is also based on a permit system in Europe and Japan.

However, though the regulatory approach finds wide application in making of the environmental policy, the approach is at the very core faulty and inefficient. This is because in this approach the same set of emission rules applies to all emitters of a specific pollutant. The policy maker plans the economic subsystems by using a general approach, and thus he is not able to take into account particular differences. Thus this approach is inefficient. Now let us take into consideration the situation of two industries A and B which are emitters of the same kind of pollutants and as a result they are being subjected to the same set of rules in that regulatory approach. Now to reduce the pollution levels by say X percent the abatement cost which will be incurred by the industries will be different, because though the industries may be producing the same kind of pollution, there will be specific differences between them. Thus, it is inefficient for the industry which incurs a higher abatement cost while undergoing the abatement procedure. Moreover, inefficiency means that the resources are being wasted, and thus the opportunity costs will be too high. Since the costs of environmental policy will have an effect on the target level, inefficient

Published in Articles section of www.manupatra.com

abatement implies less environmental quality. Therefore, the regulatory approach reduces the chances for effective environmental policy.[14]

One of the major problems faced as a result of the regulatory approach which is used by several countries, is that the regulatory approach views economy as a static entity. Thus as a result of this, since there is a pre-decided level of pollution and since a proportionate amount of permits are issued, once those permits run out the new industries cannot come up and moreover the dynamic industries which are existing cannot expand. We have already seen that this method gives protection to industries which are pre-existing and this can be inefficient because such old existing industries have old technology compared to the new industries as well as the dynamic industries and thus the utilisation of the resources will be less and inefficient and as a result the opportunity cost will be high. Closing off of a region to a newcomer reduces mobility and implies efficiency losses. All these phenomena reduce productivity or result in a slowdown of increase in productivity. Thus the regulatory process as already proved is inefficient. It is also disadvantageous to labour as the new industries are not able to locate in the region although they may provide interesting and improved employment opportunities.

INTERNATIONAL TRADE

Environmental regulations cost polluters money. Firms subject to tighter environmental regulations will incur higher costs than firms subject to weaker or non-existent environmental regulations. Thus, we can infer that economic theory would indicate that in the countries where there are relaxed environmental regulations there would be a concentration of polluting industries. Every industry would naturally (due to reasons of apparent economic efficiency) like to shift to a country where there are less environmental regulations because it would enable the industry to produce goods without investing a substantial amount of capital on either using cleaner technology or paying penalty for the pollution generated by those industries. Thus, this leads to the development of areas/countries which have only pollution-intensive industries due to its weak environmental regulations and areas/countries which specialize in clean industries due to its strong environmental regulations. Moreover, if we have significant amount of such specialization, the establishment of free trade areas would be of major concern to the environmental groups. In fact, the environment was a major issue in the debate in the United States over the North American Free Trade Agreement.[15] The issue was whether a free trade zone involving Canada, the United States, and Mexico, would result in a significant number of Canadian and U.S. Industries moving to Mexico, seeking out weaker environmental regulations. Similarly, should Eastern Europe join European Union, will firms migrate to the East where environmental regulations may be more relaxed? Furthermore, tight environmental regulations tend to drive the industries

14 Siebert, Horst, and Heinz Nixdolf. *Economics of the Environment*. 7th ed. New York: Springer, 2008. 133.

15 *Supra*. n at 6, pg. 254.

Published in Articles section of www.manupatra.com

elsewhere to places where they can find more relaxed regulations and weak government monitoring of the industries present.

Thus we can see that certain countries/ specific areas specialize in certain industries, depending upon the assimilative capacity of the area. One of the primary models of explaining international specialisation is the Heckscher-Ohlin model of International Trade.[16] The Heckscher–Ohlin Theorem states that countries export those commodities which require, for their production, relatively intensive use of those productive factors found locally in relative abundance. The twin concepts of relative factor intensity and relative factor abundance are most easily defined in the small dimensional context in which the basic theory is usually developed. Thus there are two basic things that determine the Environmental regulations of a particular country, first being the assimilative capacity of the area. For example a country with stringent environmental regulations will have clean industries, while a country having weak environmental regulations like pollution intensive industries, a country with plentiful labour will specialize in labour-intensive markets and so on. Second aspect on which the environmental regulations of a region depends on is the income of individuals[17] residing in the region. The more the income of the residents, the more stringent are the laws. This has been illustrated before.

One problem with finding evidence that differential environmental regulations influence International Trade is that the effect is probably small. Because of this, some authors have focused on the margin most likely to be sensitive to the effect: international flows of capital.[18] The logic is that, the nature of the capital stock in any country and thus that country's specialisation in a particular industry, is the result of investments made over many decades. Movement of capital from one country to another for investing in the industry of the other country is called Foreign Direct Investment (FDI).[19] If lax environmental regulations tend to attract dirty industries, then we should expect to see FDI for dirty industries concentrated in countries with lax environmental regulations, *ceteris paribus*.

Now we shall discuss one of the most pertinent problems that arises as a result of the creation of pollution havens: that is the problem of transfer mispricing.[20] It is possible that companies that need to purchase carbon credits to offset their emissions may also have investments in derivatives trading and in businesses responsible for generating carbon credits. This raises the potential for those related companies to engage in transfer mispricing. Transfer mispricing, also known as transfer pricing manipulation, refers to trades between two related parties at artificial prices for the purposes of tax avoidance.[21]

If two unrelated companies trade with each other, it is generally accepted that they will deal with each at the market price for the transaction because it is the

16 "Heckscher–Ohlin Theory | Online References | Cyclopaedia.net." Accessed October 11, 2014.
17 *Supra*. n at 6, pg. 256.
18 *Supra*. n at 5.
19 *Ibid*.
20 "Transfer Pricing - Tax Justice Network." Tax Justice Network. Accessed October 11, 2014.
21 *Supra*. n at 5.

Published in Articles section of www.manupatra.com

product of genuine negotiation in a market. But when two related companies – that is, a parent company and a subsidiary, or two subsidiaries controlled by a common parent – trade with each other, they may artificially distort the price at which the trade is recorded, to minimise the overall tax bill. This might, for example, help the company record as much of its profit as possible in a tax haven with low or zero taxes.[22]

In the case of the carbon market, for example, a company (the "parent company") may need to purchase carbon credits to offset its own emissions. It may invest in an offset project in Africa to generate those carbon credits. But first it establishes a trading company in a tax haven (with zero taxes). The subsidiary company in Africa sells its carbon credits (at a very low price) to the trading company, who then sells those carbon credits (at a very high price) to the parent company. In this case the subsidiary company in Africa receives an artificially low price for the carbon credits, resulting in low profits – and consequently an artificially low tax bill in Africa. The company in the tax haven sells the credits at a high price – artificially transferring all of the profits there, but being a tax haven, no tax is paid.[23]

Thus, as a result of the above phenomenon, we see that there has been complete avoidance of payment of taxes in the country in which the offset project was set-up, the polluting industry in the main country can produce unlimited pollution but it can save itself from any adverse effect, because they have an unlimited source of production of carbon-credits which will meet their requirement. Thus, we see that such tax havens or pollution havens which are developed at specific countries due to their lax environmental regulations have an extremely negative effect upon the environment. We have to remember that the basic reason why the environmental regulations were enacted in the first place was to protect the environment from degradation and overall welfare of the society as a whole. Thus, when one industry can generate immense amount of pollution, not caring about the carbon-cap because it has an unlimited supply of carbon credits, it is extremely detrimental to the overall environment. This leads to the loss of the very reason for which such carbon caps were incorporated in the system in the very first place. If the environmental pollution is not reduced and the polluting industries continue to carry out their detrimental procedures, then the institution of carbon credits and the carbon-related regulations is of no use. Moreover from the economic perspective, such transfer mispricing also leads to tax fraud in the country (say Africa, in the example stated above) in which the set off industry is set up to produce carbon credit, because the set off industry sold off the carbon credits at an extremely low price and as a result the regulatory authorities could not charge the normal taxes or charged zero taxes for such transaction because of the extremely low financial amount involved in the transaction. Thus we see that the economic benefit that the country (Africa in the above example) should be receiving from the taxes are lost because of the fraud. Thus, this situation is not Pareto efficient because Pareto efficiency requires that one

22 *Ibid.*
23 *Ibid.*

Published in Articles section of www.manupatra.com

person should be better off (in his own estimation) and no other person may be made worse off (again in his own estimation). Thus, clearly from the above situation we can see that the polluting industry is made better off as a result of the entire set up of the tax haven/ or pollution haven (this is apparent from the first glance), however, the country in which the set off industry is situated is made worse off due to the tax fraud and the overall society is made worse off as a result of the environmental degradation. Thus, we see that the set up as a whole is inefficient on the whole from both the economic as well as the environmental prospective.

TAX FRAUD IN INTERNATIONAL TRADE

Carbon trading in the European Union has already been hit by a particular form of tax fraud involving the theft of Value Added Tax ("VAT") (also known as carousel fraud or Missing Trader Intra-Community fraud).[24] This type of fraud exploits how VAT is treated in multi-jurisdictional trading in Europe. In this type of fraud the criminal utilises the fact that VAT is not charged when a good is traded across the boundary between member states of the European Union, thus a good say from Country A is traded to a Country B and since the countries are member states of the European Union such a transfer of that good is VAT-free. Now the person 1 who obtained that good by means of the trans-boundary trade will sell it to another person within his country and such a sale will include both sale price and VAT. Now the person 2 who obtains this good is supposed to pass on the VAT to the relevant government to whom this is owed, however in the situation where there is a tax fraud such a person usually disappears without paying the relevant tax and subsequently the goods and the VAT may then be sold on through a number of companies, passing across a number of borders, with each additional layer of transactions making it more difficult to trace the link between the final VAT that is owed to the relevant government authority and the original importer, who by this time has disappeared without paying the tax.

In 2009 authorities began to observe high volumes of trade on France's BlueNext carbon exchange. Subsequent investigations revealed the sudden spike in trading was the result of highly organised criminals trading large volumes of carbon credits as part of a carousel fraud scheme. The European police agency, Europol, subsequently estimated that up to 90% of all carbon trading in some countries was a result of these fraudulent activities.[25] This fraud was estimated to have resulted in losses to several governments of around 5 billion euros in just over 18 months.

Morevover, in the case of *R* v. *Dosanjh, Chahal and Gill* [2012], in June 2012, at Southwark Crown Court in the United Kingdom, three defendants were found guilty of carbon trading carousel fraud and jailed for a combined total of 35 years.

24 *Supra* at 5.

25 Mason, Rowena. "Copenhagen Climate Summit: Carbon Trading Fraudsters in Europe Pocket €5bn." THE TELEGRAPH, December 10, 2009. http://www.telegraph.co.uk/earth/copenhagen-climate-change-confe/6778003/Copenhagen-climate-summit-Carbon-trading-fraudsters-in-Europe-pocket-5bn.html.

Published in Articles section of www.manupatra.com

The three VAT fraudsters had established a number of bogus companies to import the carbon credits into the UK, and when those credits were sold on the original companies were dissolved. The credits were sold on again between three further 'buffer' companies – also run by the gang – to make the trading chain appear legitimate, before the credits were finally sold to legitimate companies, charging VAT which was never paid back to the government. During the 69 days of trading, the total turnover was 276 million euros, of which 41 million Euros was due in VAT. The trades were made in a matter of minutes via a computer system, and the stolen VAT was transferred to offshore bank accounts in the United Arab Emirates to 'clean' the stolen cash which the gang then spent.[26]

Thus, we see as a result of such tax fraud by not paying the VAT to the required authorities there has been an immense amount of economic loss to the respective countries and as a result we see that such processes are basically inefficient and cause widespread economic loss. Now, if we look into the background as to why frauds in general take place, we will see that while committing such offences the major thought process of such violators is that even if they get caught the penalty which they will have to pay or the loss which is suffered is less than the total profit that is generated as a result of this violation. Now, this is efficient according to Potential Pareto Improvements or Kaldor-Hicks efficiency.[27] Now this theory of efficiency was developed by economists who wanted to overcome the restriction which was laid down by the Pareto criterion, which is, to achieve Pareto efficiency, one person must be made better off while no other person can be made worse-off. However, economists wanted to develop a more practical approach and as a result they developed the Potential Pareto Improvements or Kaldor-Hicks efficiency which said that in a situation there can be both persons, one who has obtained a profit and another who has suffered a loss however the person who has obtained a gain must be better off than the person who has suffered a loss, i.e., the gainer's gain must be more than the loser's loss and as a result even after compensating the loser the gainer must have surplus left as a result of which such a situation is efficient.

However, the situation which is being dealt with in the present scenario is not efficient according to Kaldor-Hicks efficiency. Here, the loss of the loser is greater than the gainer's gain, which is very evident from the reports that specify that around 90% of trade in carbon credit is as result of fraudulent activities (in certain countries) and moreover there is no instance of payment of compensation by the violator to the person who has suffered the loss as these people do not pay compensation for the VAT fraud as it is extremely difficult to locate such violators due to the presence of a number of transactions and foreign exchanges and the

26 Alllen, Emily. "Fraud Mastermind Cheated Taxpayer out of £39million in Just 69 DAYS and Spent the Money on £1million Home and a Rolls Royce Is Jailed for 15 Years." *Mail Online*, June 19, 2012. Accessed October 11, 2014. http://www.dailymail.co.uk/news/article-2161520/Fraud-mastermind-cheated-taxpayer-39million-just-69-DAYS-spent-money-1million-home-Rolls-Royce-jailed-15-years.html#ixzz29O6SfXtp.

27 Cooter, Robert, and Thomas Ulen. *Law and Economics*. 4th ed. Delhi: Pearson, 2004. 48.

Published in Articles section of www.manupatra.com

relaxation in regulation and cross checking in these foreign exchanges. As a result such a situation is not efficient and is highly detrimental to the global economy.

CONCLUSION AND SUGGESTIONS

Thus, we see that when we go on to analyse International Trade or the provisions on the basis of which regulations (legislations) are made in countries, there are severe defaults in the practical application of theoretically successful provisions. As illustrated above, we have already seen that income and assimilative capacities are two major aspects which influence the making of laws. Economic benefits have been the only drive which has influenced the countries and Industries. It is true the developing or under-developed countries cannot afford to implement stringent environmental regulations because they need the development of industries for their economic development, but this factor has been exploited by several factories, and thus has resulted in areas specializing in polluting industries giving rise to pollution havens. Thus we see that there is an overall detrimental effect on the environment, because on one hand there are countries, which have implemented strong environmental laws to improve the quality of the environment and on the other hand there are these countries with weak environmental regulations which are polluting the environment. Since, environment is a public good and can be used in equal amounts by all the countries, individuals or groups can take the position of a free rider not contributing to the cost of environmental quality. Constitutional mechanisms have to be developed which ensure that free-rider behaviour in evaluating environmental quality is reduced.

Now, coming to the concept of whether subsidies or penalties should be made applicable for such environmental pollution, although we have seen that the Coase theorem says that penalties and subsidies have the same effect on efficiency, practical application has shown that providing subsidies is inefficient. Thus, we can say that practical application of certain laws should be reconsidered in terms of its environmental implication and if such laws are not consistent with the environment/or have a detrimental impact on the environment, such laws should be removed because Environmental laws should not only be economically efficient but also promote the betterment of the environment (in its practical application) because that is the basic reason of the institution of such laws.

We have already seen the problems associated with International Trade. Thus, let us see the provisions which may help to improve the situation. One of such provisions is the implementation of sanctions against such polluting countries. Trade is one of the several aspects that controls international relation, however, there are several other aspects that govern international relations, like the fact that countries deal with each other on a multitude of issues, not only economic but political and cultural as well. It may be that the victim country has little ability to coerce the polluter to reduce emissions; however, the victim country may have a great deal of power on a totally unrelated issue, such as involvement in a joint defence treaty. By withholding agreement on the joint defence treaty, the victim

Published in Articles section of www.manupatra.com

country may be able to force the polluter to reduce pollution levels. In essence, the victim country can threaten sanction of one sort or another against the polluter country if it does not adopt pollution regulations to achieve the efficient amount of pollution. Another means of controlling the pollution situation is by having a stringent multinational Agreement which involves as many countries as possible so that the deviance from the agreement is not in the scope of events, because it is only when majority of the countries are not subject to a uniform law that problems like pollution haven and tax haven arise. The basic problem in designing an effective international environmental agreement is that there is no super-national organization to enforce such an agreement. In effect, the agreement must not only be self-enforcing but it must be sufficiently appealing to all parties involved for them to agree to it in the first place. Thus there are three desirable characteristics of an international environmental agreement:

(1) the agreement should be self-enforcing;

(2) each country should be better off as part of the agreement than not; and

(3) pollution levels obtained by the agreement should be Pareto improvement over the status quo and, ideally, a Pareto optimum.

It is easy to contemplate an international agreement that meets the first two of these three criteria; it is quite another thing to design an agreement that meets all three conditions.[28]

There are several principles which have been proved to be economically efficient and may be applied or taken into consideration when policies are made so that not only they benefit the environment but also are efficient economically. The polluter-pays principle of environmental policy is an institutional manifestation of the opportunity cost principle.[29] It can be applied once environmental targets are established. Now, this system ideally expects a global target to be established, in such a situation that no country should be left out of its ambit and as a result this principle helps in circumventing the free-rider problem because in such cases the absence of a global law or target allows certain countries to ignore the requirements of the environment while other countries are being bound by stringent environmental regulations. This principle helps in allocating the opportunity costs of the environmental protection in a reasonable way. The individual polluter has an incentive to reduce pollutants (as the pollution which will result from his actions needs to be compensated by him), the divergence or gap between the private costs and social costs are abolished and as a result of this the commodity price includes all the factors, i.e. the price of production of the good and the environmental cost. This enables the consumer to see the true price of the commodity and thus the consumer will decide the consumption accordingly. The polluter-pays principle can take many forms such as emission taxes, compensation procedures as in Japan's environmental

28 *Supra*. n at 6 pg. 264.
29 *Supra*. note at 14, pg. 164.

Published in Articles section of www.manupatra.com

policy, or liability rules.[30] The polluter-pays approach is the documentation of the more general question of an appropriate institutional setting for allocating the opportunity costs of the environmental protection to the subsystems. It may be interpreted as a solution to an incentive problem.

However, there are several constraints to the polluter pays principle,[31] the primary one being the identification of the polluter which is the most important part for the execution of this principal. Second constraint is the vote-maximisation intention of the policy maker. Thus, this one constraint will become a disincentive for the policy maker from making policies which will induce the polluters to pay for the pollution caused by them. And, the third constraint is the fact that if the country or region is located in an upstream or upwind position then they will have an incentive in asking for compensation or application of the pollutee-pays principle rather than application of the polluter-pays principle.

Now, this brings us to the concept of pollutee-pays principle[32] which is the counter principle of the polluter pays principle and this principle talks about the notion in which the pollutee has to compensate the polluter in order to induce him to reduce or avoid pollution. This arises from the context of the Coase-theorem. This principle may be applied in the case of global media when high income countries place a higher value on preventing climate, risks and compensate the lower income countries through payments in order to induce them to avoid pollutants.

Another important consideration which needs to be taken care of is the *Principle of Interdependence*[33] which talks about how all the environmental systems are interdependent and represent a complex system or network of interaction. It flows naturally that since the environmental systems means the interdependence of pollutants due to the fact that pollutants are linked through environmental systems and diffusion between them. For example, the pollutants ambient in the air can be deposited into water systems, and pollutants ambient in the river and other water bodies reach the air by means of the process of evaporation. Pollutants may also be linked by means of emission technology.

Environmental policy must take these interdependencies between environmental media, between abatement, emission and production technologies, and between pollutants into account. If environmental policy addresses itself to only a particular media, a particular pollutant, or a particular abatement or production technology, it is likely to fail in the long run. Very quickly, new problems will pop up. Consequently, environmental policy has to be integrative and encompass all environmental media and pollutants.

30 *Ibid.* at pg. 165.
31 *Ibid.*
32 *Ibid.* at pg. 166.
33 *Ibid.* at pg. 167.

Published in Articles section of www.manupatra.com

PAPER • **OPEN ACCESS**

Analysis of Environmental Law Enforcement Mechanism Based on Economic Principle

To cite this article: Hongjun Cao *et al* 2017 *IOP Conf. Ser.: Earth Environ. Sci.* **94** 012188

View the article online for updates and enhancements.

You may also like

- Environmental law enforcement in forestry crime: A disjunction between ideality and reality
S Muchtar and A Yunus
- Future environmental and agricultural impacts of Brazil's Forest Code
Aline C Soterroni, Aline Mosnier, Alexandre X Y Carvalho et al.
- The importance of national political context to the impacts of international conservation aid: evidence from the W National Parks of Benin and Niger
Daniel C Miller, Michael Minn and Brice Sinsin

This content was downloaded from IP address 37.101.222.112 on 01/11/2022 at 18:08

Analysis of Environmental Law Enforcement Mechanism Based on Economic Principle

Cao Hongjun[1, 2, a], Shao Haohao[2, b], Cai Xuesen[1, c]

[1]School of Management, Ocean University of China, Qingdao, China

[2]School of Environmental Science and Engineering, Ocean University of China, Qingdao, China

E-mail: [a] hongjunsd@163.com, [b]shaohh@sina.com, [c] 963401293@qq.com,

Abstract: Strengthening and improving the environmental law enforcement mechanism is an important way to protect the ecological environment. This paper is based on economical principles, we did analysis of the marginal management costs by using Pigou means and the marginal transaction costs by using Coase means vary with the quantity growth of pollutant discharge Enterprises. We analyzed all this information, then we got the conclusion as follows. In the process of strengthening the environmental law enforcement mechanism, firstly, we should fully mobilize all aspects of environmental law enforcement, such as legislative bodies and law enforcement agencies, public welfare organizations, television, newspapers , enterprises, people and so on, they need to form a reasonable and organic structure system; then we should use various management means, such as government regulation, legal sanctions, fines, persuasion and denounce, they also need to form an organic structural system.

1. Introduction

With the rapid development of social economy in China, the function of the government has changed dramatically. From mainly through administrative means to mainly through economic means. Facing the increasingly severe ecological problems and the transformation of government functions, it is urgent to strengthen and improve the environmental law enforcement mechanism.

In recent years, many scholars have conducted a more in-depth study of the environmental law enforcement mechanism. By making the analysis of the application of economic instruments for environmental management in China, Mo Chuangrong (2005) pointed out three conditions for the implementation of economic instruments: legal guarantee, the effective market mechanism, the rationalization of government behavior; meanwhile he put forward seven suggestions to solve the issues of environmental management development. Xu Zhonglin(2013)pointed out that the three aspects are complementary to each other, forming a system of environmental law enforcement mechanisms, and jointly playing the role of environmental law enforcement. From the angle of "Ecologicalized law enforcement", Li Ainian(2016) pointed out some suggestions, such as perfecting the mechanism of government accountability, strengthening the power of administrative enforcement of law, strengthening the construction of law enforcement capability , establishing diversified law enforcement methods and so on. Sjöberg E.（2016）pointed out that politicians do not only affect environmental policy, but also that for a given policy, they can affect the outcome through implementation and enforcement. Stretesky (2017) provides strong empirical support for the hypothesis that the global environmental enforcement culture is shaped by competitive neoliberal tendencies. On the basis of previous studies, this paper explores and studies the environmental law enforcement mechanism of environmental

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

supervision departments in the process of increasing quantity of sewage enterprises from the perspective of economic principles, and provides theoretical references for the relevant departments.

2. Application of economic principle in environmental law enforcement mechanism

2.1 The Application of Pigou Means in Environmental Enforcement

The most important application of Pigou means in environmental enforcement is to levy shelter tax. Implementing negative incentives for the enterprises that produce negative external effects, such as environmental problems. Otherwise, implementing positive incentive, which is grant subsidies, for the enterprises that produce positive external effects. In order to obtain balance by applying Pigou means. Assuming that the remaining conditions are certain, the choice of optimal combination of environmental governance methods depends on the marginal transaction costs and marginal management costs in the process of environmental governance. In the application of Coase means in environmental protection, marginal transaction cost is the amount of transaction cost increase among enterprises by adding one more environmental pollution enterprise. In the application of Pigou means in environmental protection, marginal costs of management is the amount of total cost of the governance environment increase by adding one more environmental pollution enterprise.

2.2 Application of Coase Means in Environmental Enforcement

Based on the general sense of Coase means, if the property right is cleared, the transaction costs approach to zero. The allocation of resources will ultimately achieve to Pareto Optimality, achieving efficient market equilibrium regardless of the side of property. But in reality, transaction costs can't be zero. The more companies pollute, the greater the transaction costs will be. In the process of environmental governance, Coase means represents a negotiable or tradable license system. Pigou means focus on the government regulation while Coase means pays more attention to the role of market mechanism. Based on Kos theorem, the most effective and reasonable solution appears when the property right is clear, nointervention among different interests and resolve disputes by voluntary negotiation.

3. Economic Analysis of Environmental Law Enforcement Mechanism

3.1 Optimal choice model of economic measures of environmental management

Assuming that the technical level of the firm is constant, the amount of emissions is at the same level and there is no significant difference in the environmental benefits between the enterprises, regardless of the transaction costs of the beneficiaries that may bring about the efficiency of the loss and the use of monopoly for high returns of rent-seeking behavior, the choice of the best combination of environmental management methods depends on the choice of environmental governance in the process of marginal transaction costs and marginal management costs. If using the Pigou means to manage the environment, one more environment contaminated enterprise will increase how much the total cost of governance environment, the increase quantity of the total cost is the marginal management cost. Marginal management cost include the growth of management costs associated with the production of environmental-related institutions, the increase in monitoring costs of monitoring departments and the imposition or payment of taxes or subsidies; a rise in costs resulting from the polluting enterprises. If using the Coase means to manage the environment, how much more a sewerage company will add to the transaction costs, the increase in transaction costs is the marginal transaction costs. On the basis of government management costs without regard to government control, under the market mechanism, the interests of the parties through the use of voluntary consultation, the transaction costs include the cost of victims to understand the sewage business sewage situation, the cost of victim clear their own degree of damage and other information and the cost of the negotiation before. When using a sewage permit transaction, the transaction costs include the cost of finding the transaction object and the cost of the contract. The combination of the ways of environmental management can be optimized according to

Figure 3-1:

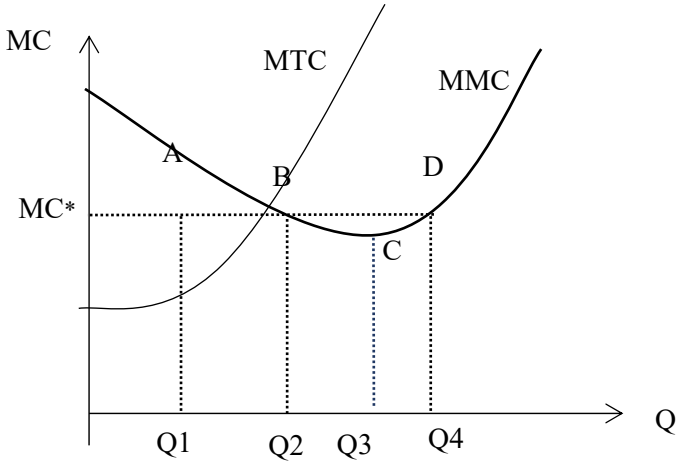


Figure 3-1. Optimal choice model of economic measures of environmental management[1]

As shown in Fig. 3-1, the horizontal axis represents the quantity of pollutant discharge Enterprises, represented by the letter "Q"; the vertical axis represents the marginal cost expressed by "MC". MMC curve represents the marginal management costs by using Pigou means vary with the quantity growth of pollutant discharge Enterprises，MTC curve represents the marginal transaction costs by using Coase means vary with the quantity growth of pollutant discharge Enterprises. As can be shown from the figure，the change of MMC curve was shown as "U" shape. In the process of the quantity of pollutant discharge Enterprises grow up from 0 to Q3, the curve leans toward the lower right until reduced to the lowest point "C"，This shows that marginal management costs continue to decline as the quantity of polluters increased. Then the MMC curve rises gradually from the point Q3, marginal management costs continue to rise as the quantity of polluters increased. Meanwhile, the MTC curve has been tilted upward, and the margin of inclination is increasing. That is, marginal transaction costs continue to rise with the increase of the number of enterprises, and the rate of increase is increasing faster and faster. This can be represented as a formula：

$$\frac{\partial^2 MTC}{\partial Q^2} \geq 0 \quad (1) \qquad \frac{\partial MTC}{\partial Q} \geq 0 \quad (2)$$

There is a critical point B in Fig.3-1, and the coordinate of this point is (Q2, MC*). That is to say, when the quantity of enterprises is Q2, the corresponding marginal cost is MC*. The B point is the intersection of the MMC curve and the MTC curve. At this special point MMC = MTC, it is the same to choose the Pigou means to control the environment, or to choose the Coase means to control the environment. The number of pollutant discharge enterprises Q2 corresponding to B point is the number of critical polluting enterprises. When Q is between 0 and Q2，the marginal transaction cost MTC is less than the marginal management cost MMC. At this time, Coase means should be chosen to govern the environment. If further refined: (1) When Q is between 0 and Q1，the number of pollutant discharge enterprises is very small，and the whole trading market is imperfect. In this case, the most effective way is to negotiate voluntarily among stakeholders.(2) When Q is between Q1 and Q2，the trading market is relatively perfect. In this case, the choice of emission permits trading is effective. In addition, Q1 is not a strict dividing point. The boundaries between voluntary negotiations and licensing are not clear.

When the number of pollutant discharge enterprises exceeds Q2，marginal transaction costs MTC

[1]Ma Zhong. Introduction to environment and Natural Economics [M]. Beijing：Higher Education Press, 2006

exceed marginal management costs MMC. In this case, Pigou means should be chosen in environmental governance. As can be shown from the figure, the change of MMC curve was shown as "U" shape. After dropping to the lowest point C, the curve began to pick up sharply. After the point C, choosing which method for environmental governance is an important matter. We should not only to consider the size relationship of MTC and MMC, but also to compare the government's management costs and control costs, when using Pigou means. As can be shown from Figure3-1, When Q is between Q3 and Q4, marginal cost is less than MC*. In this case, the use of Pigou means of environmental governance can be effective. But when the number of enterprises is greater than Q4，the marginal management costs will rise sharply and disputes among subjects of interest will soar. In this case, the use of economic incentives to control the environment will be ineffective. The best choice at this moment is adopting government coercive measures to carry out environmental governance.

3.2 Analysis of market mechanism improvement and government efficiency improvement

Suppose the MMC Curve is fixed, with the development of the marketing mechanism, as shown in Fig. 3-2, the MTC Curve will move rightly to the MTC' curve. Then the intersections of marginal management cost curve MTC and marginal transaction cost curve will move rightly from Point B to Point Y, the number of the pollutant discharge enterprises will grow up from Q2to Qy, adopted by Pigou means or Coase means. Suppose the MTC curve is fixed, with the improvement of the government's working efficiency and the environmental enforcement capacity, the MMC curve will go down to the MMC' curve, the intersection of the marginal cost curve and marginal transaction cost curve will move from point B to point Y' ,the number of the pollutant discharge enterprises will decline from Q2to Qy', adopted by Pigou means or Coase means, which means with the improvement of the government's working efficiency, the section of environmental management will be narrowed by using Coase means, which will be declined from Q2to Qy'. It also means the the section of environmental management will be widen by using Pigou means.

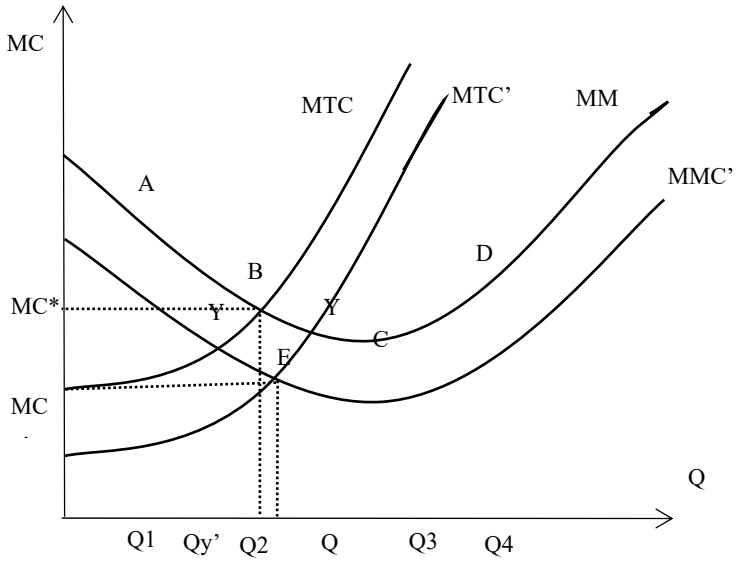


Figure 3-2. Reasonable structure model of environmental control measures[1]

If both the market mechanism and government's working efficiency are continuously improving, environmental awareness of citizens is also enhanced, then the MMC curve and the MTC curve will

[1]Ma Zhong. Introduction to environment and Natural Economics [M]. Beijing：Higher Education Press, 2006

move to MMC' and MTC'. Thus the intersection in the Fig. 3-2 will move from point B to point E. It can be shown ,there is little different between the quantity of pollutant discharge enterprises related to point B and point E .which means the influence of the Pigou means and Coase means on section of environmental management is the same. The marginal cost corresponding to point B to point E will decline from MC* to MC** during this decline process. The social force has also plays an important role. Therefore, in terms of the consideration of the settlement mechanism of environmental problems, apart from environment itself, the system background, political environment and social security should also be taken into consideration.

4. Conclusions and Suggestions

Based on the above analysis, we got the conclusion: In the process of strengthening the environmental law enforcement mechanism, on one hand, we should fully mobilize all aspects of environmental law enforcement, such as legislative bodies and law enforcement agencies, public welfare organizations, television, newspapers , enterprises, people and so on, all of them should play their own role to make up a reasonable and organic structure system；on the other hand, various management means, such as government regulation, legal sanctions, fines, persuasion, denounce, they also need to form an organic structural system.If regulation and economic incentives failed, in order not to cause chaos, it is necessary to give full play to the role of public forces in society, and put this force in a fundamental position. Apart from relying on the automatic regulation of government control by the "Pigou means" and the market mechanism by the "Coase means", we should also constantly improve people's environmental awareness and improve the informal institutions and social mechanism of environmental protection.

Above all we put forward the following suggestions: Continue to improve the market mechanism; Improve the efficiency and law enforcement capacity of environmental management department; Enhance public awareness of environmental protection; Improve environmental protection systems and measures; Establish and improve environmental guidance and encouragement mechanism; When necessary, the government should adopt compulsory means to carry out environmental management.

Acknowledgment

Shandong province science and technology development plan project "Shandong province environmental law enforcement system construction research" (Approval number: 2014GGH222001).

References

[1] Stretesky P. B., Long M. A. Lynch M. J.. Trends in the Formation of Environmental Enforcement International Non-Governmental Organizations, 1950–2010[J]. 2017.
[2] Sjöberg E.. An empirical study of federal law versus Local Environmental Enforcement[J]. Journal of Environmental Economics & Management, 2016, 76:14-31.
[4] Xu Zhonglin, Fang Xiaoan. Analysis of Environmental Law Enforcement Mechanism [J]. Journal of Jiangxi University of Science and Technology, 2013(4):14-17.
[5] Li Ainian, Liu Ao. Environment Law Enforcement in Ecological Approach:The Innovation of Law Enforcement Mechanism for the Ecological Civilization Construction [J]. Journal of Social Science of Hunan Normal University, 2016, 45(3):80-88.
[6] Mo Chuangrong, Zhang Shijia, Chen Xingeng. Application and Countermeasures of Economic Instruments for Environmental Management in China [J]. Environmental Protection Science, 2005, 31(5):26-29.
[7] Ma Zhong. Introduction to Environment and Natural Economics [M]. Beijing：Higher Education Press, 1999(6)

PAPER • **OPEN ACCESS**

A study on the relationship of environmental regulations and economic performances

To cite this article: Jia Nisha and Shen Chen 2017 *IOP Conf. Ser.: Earth Environ. Sci.* **94** 012035

View the article online for updates and enhancements.

You may also like

- The impact of fiscal decentralization on economic performance in Indonesia
F Setiawan and A F Aritenang
- The persistence of flexible coal in a deeply decarbonizing energy system
Yongbin Ding, Mingquan Li, Ahmed Abdulla et al.
- Influence of heterogeneous environmental regulation policies on the strategy of pollutant discharge for enterprise: an evolutionary game approach
Zhiqiang Zhou, Huanhuan Feng, Huilin Wang et al.

This content was downloaded from IP address 37.101.222.112 on 01/11/2022 at 18:07

A study on the relationship of environmental regulations and economic performances

NISHA Jia[1,2], and CHEN Shen[2]

[1] International Business School, Shaanxi Normal University, Xi'an 710119, China
[2] China（Xi'an）Institute for Silk Road Research, Xi'an 710100, China
[3] College of Economics and Management, Zhejiang SCI-Tech University, Hangzhou 310018, China

jianisha@snnu.edu.cn, littleangle6@sina.com

Abstract: This paper analyzes the mechanism transmission of environmental regulation affecting on economic performance from a new perspective. It shows that environmental regulations affect country's economic performance through direct transmission and indirect transmission. Direct transmission means that environmental regulations affect economic performance from influencing corporation's cost and revenue, and indirect transmission is that environmental regulations affect economic performance through leading to the increased effort of corporations in improving the quality of environment and building the image of fulfilling environmental responsibility.

1. Introduction

The economic development processes of many countries, especially some developed countries, seem to prove that economic growth mode is at the expense of environment. Due to environment resources' nature of public goods, environment problem's negative externality, and opportunism of microeconomic units, environment problem is hard to be solved with market mechanism, so environment regulation is deemed as an important tool for correcting market failure. Can current environmental regulation make economic growth and environmental protection achieve a kind of balance?

According to existing literature, there are four views on the relationship between environment regulation strength and economic development level: (1) "the former has a negative impact on the latter". This view's theoretical base is "compliance cost theory". Represented by Gollop and Roberts(1983), Barbera and McConnell(1990), Walley and Whitehead(1994) , the scholars holding this view think environment regulation causes new cost factors to enterprises, which are called "compliance cost". The costs will bring about "extra" burden to enterprises', affect output and profit level, and weaken enterprises'' competitiveness. Besides, corporate pollution control expenditure may occupy other production investment. When resources are limited, environment regulation will indirectly increase enterprises' opportunity costs. (2) "the former has a positive impact on the latter". Its theoretical base is "Porter Hypothesis" (Porter and Van der Linde, 1995; Porter, 1996). The scholars advocating this view think choosing proper environmental regulation can drive enterprises to innovate, induce the "innovation offset" effect between production process and products (namely technological innovation offset and product innovation offset), and thus reduce enterprises' costs, further bring about technology diffusion effect, improve enterprises' productivity, achieve the optimization and upgrading of enterprises'' own structure and industrial structure, and finally improve

Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.
Published under licence by IOP Publishing Ltd

their market competitiveness. However, "Porter Hypothesis" has two preconditions: the first is that resources configuration, technology, consumption demand and other conditions are variable dynamic model; the second is that environment regulation tool must be "properly designed". (3) Uncertain "mixed view". Those holding the "mixed view" think the results of mixed effects are uncertain or the results are inconsistent for different market subjects. (4) Environmental regulation's influence on economic growth has inflection point. Some domestic scholars make empirical test on the relationship between environment regulation and economic development respectively from the levels of state, region, industry and enterprise, and find the relationship between the two shows a U-shaped curve (Fu Jingyan and Li Lisha 2010; Xiong Yan, 2012).

Scholars have analyzed the relationship between environmental regulation and economic growth from different points of view, providing reference for the research of this paper. However, there is little literature considering the direct transmission mechanism and indirect transmission mechanism. This paper attempts to systemic analyze the mechanism transmission environmental regulation affecting on economic performance.

2. Analysis on Theoretical Mechanism

According to existing theories, environmental regulation directly affects corporate performance. The sum of net effect of emission controlling enterprises form the net efforts on department economy and finally form the net influence on regional and national economic performance. The action channels of environmental regulation's influence on enterprises have both direct and indirect transmission mechanisms (as shown in figure 1).

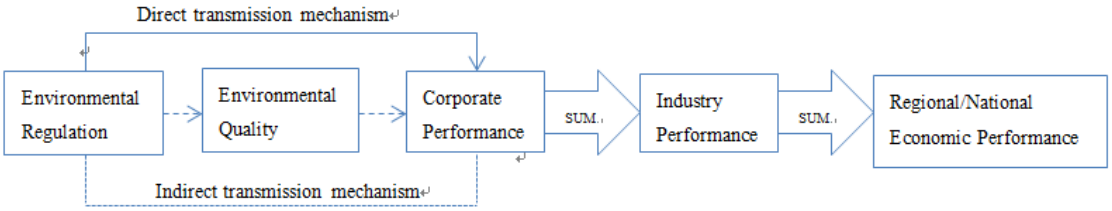


Figure 1. The Mechanism of Effect of Environmental Regulation

2.1. Direct Transmission Mechanism

Direct effect has both positive and negative sides. Positive effects include: (1) stimulate corporate technical innovation, and establish competitive advantages. As environment regulation improves, government's preferential policies for the related industries will bring about opportunities for innovation for enterprises, and drive the development of new environment protection materials and equipments and optimization of production technologies. This product (innovation) compensation effect can increase products' value or reduce production costs. Environment regulation will also drive enterprises to develop pollution control technologies to optimize its production process or waste management and improve resources' productivity. This effect is called as process (compensation effect). Product compensation and process compensation effect work together to promote enterprises to establish and keep competitive advantages through innovation. (2) Improve market access "threshold" and prevent new competitors from entering. After the implementation of environment regulation policies, existing enterprises enjoy greater cost advantages because of mature pollution treatment technologies and scale economy. The mechanism that obstructs new enterprises from entering and reduces industrial competition is called access barrier effect. Enterprises may be driven to invest on emission reduction by environment constraint and thus be subject to the decline of performance: (1) demand declines as enterprises' production and transaction costs are increased. Environment regulation raises the direct production costs such as pollutant emission tax/charge, pollutant emission license and prices of inputs and transaction costs such as pollution measuring and monitoring and negotiation with government and supplier, which is unfavorable for simple and expanded reproduction and the subsequently increasing sales price may also reduce the demand on products. (2) Occupy scarce resources and squeeze production investment. New emission reduction

equipments and R&D and use of production technology require a lot of early-stage investment and personnel training. Transition to new equipments also needs certain conversion and scrap costs. At the same time, enterprises' scarce resources, such as high-quality talents are used to improve environment quality, so production investment is occupying and resources utilization efficiency is also limited.

2.2. Indirect Transmission Mechanism

The external pressure imposed by environment regulation will drive enterprises to improve environment quality, deliver "responsibility investment" and "green investment" information to the society, and thus bring about feedback effects of different stakeholders to enterprises. The indirect transmission mechanism of environment regulation's influence on corporate economic performance works as follows: (1) promote product characterization and create new demands. As consumers attach more importance onto environment protection, green industries and the enterprises better performing their environment protection responsibilities are increasingly preferred by market. Environment regulation will drive enterprises to change the passive emission reduction to active environment treatment investment, establish distinct product characteristics, stand out from competition and thus improve products' sales and added value. (2) Improve enterprises' brand and fame and develop intangible assets. Stakeholder theory and the theory on enterprise's social responsibility reveal the importance of intangible assets such as corporate culture and fame. Improving environment quality can help enterprises build brand and fame, expand social recognition and influential power, and the appreciation of intangible assets can also improve enterprises' performance in capital market. (3) Improve stakeholders' loyalty, and reduce risk and operational costs. Enterprises' efforts on reducing environment pollution help improve the loyalty of stakeholders including consumer, suppliers and employees, reduce enterprises' risk and operational costs, for example reduce high-quality employee turnover, reduce financing difficulty and costs, and lower the price and insurance fee of raw materials etc. A good relationship with government can also reduce enterprises' compliance costs, such as tax, licensing costs, penalty and lawsuit costs etc.

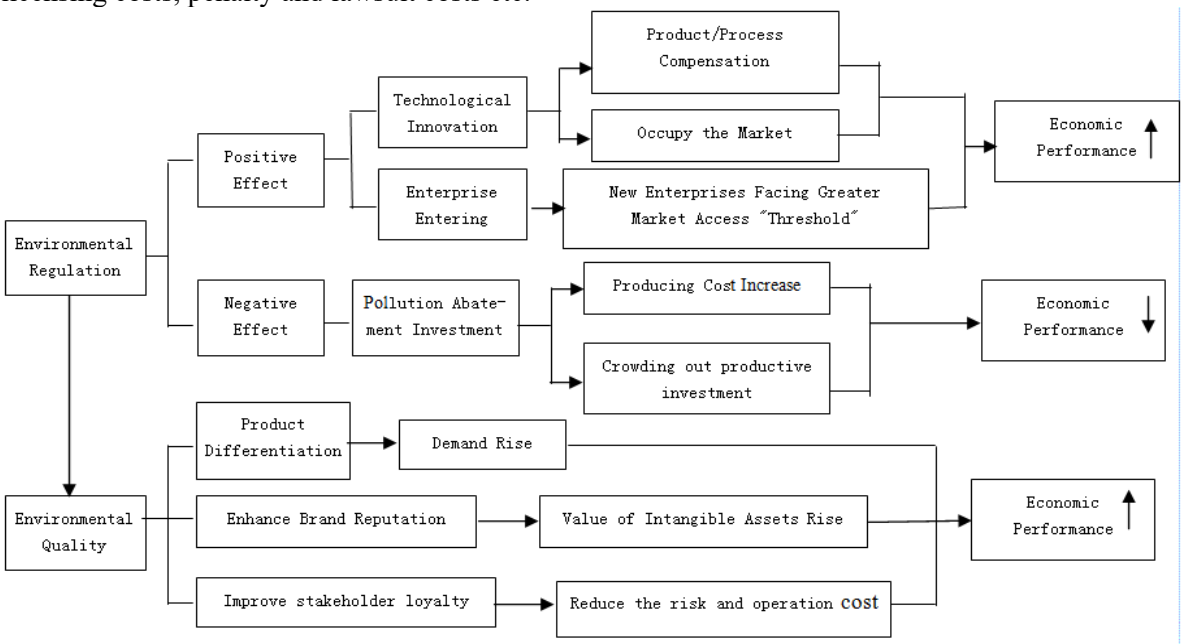


Figure2. Direct and Indirect Transmission Mechanism

3. Conclusion

In this paper, the theory analysis show that environmental regulation directly and indirectly effects the economic performance, and environment regulation's influence on corporate economic performance is quite uncertain (as shown in figure 2). Its net effect depends on the relative degree of direct and

indirect transmission mechanism influenced by multiple factors such as enterprises' characteristics, specific environment problem, policy tool, industrial structure and location characteristics and the interaction of the mechanisms.

4. Policy Suggestion

Based on the conclusions from analysis above, we draw the following policy recommendations:

First, Leaders should strengthen the legal consciousness, and make sure administration according to law strictly. We should improve the situation of legal system that "there shall be laws to abide by, everyone should abide by the law, the law must be enforced strictly, and those who violate the law must be dealt with". Advancing the reform of environmental management system and implementing the vertical administration system to make sure that the environmental protection department has an independent monitoring and administration. Defining the legal situation of the status of environmental monitoring and institutions, the governor should gradually increase level of environmental monitoring agencies and the authority of environmental enforcement team based on the actual situation.

Second, it should reduce the degree of association of the economic performance and officials promotion, and add the regional environmental benefits, the consumption of resources and energy into the evaluation system, and enhance the incentives and constraints of public welfare in the region in the establishment and implementation of a wide range of local government performance evaluation system. In the pursuit of economic performance, establishing the public service-oriented "service type" local government, so as to promote the coordinated development of regional economic and ecological.

Third, improvement of ecological environmental quality as a result of environmental regulation constraining enterprise pollution behavior can promote the improvement of corporate performance. Thus, the local government should encourage enterprises to strengthen environmental responsibility, increase investment in environmental governance and fulfill corporate social responsibility. This is conducive for us to improve both environmental quality and economy growth.

References

[1] Gollop R 1983 Environmental regulations and productivity growth: the case of fossil-fueled electric power generation *The Journal of Political Economy.* **91** 654-674.

[2] Barbera AJ and McConnell VD 1990 The impact of environmental regulations on industry productivity: direct and indirect effects *Journal of Environmental Economics and Management.* **18** 50-65.

[3] Walley N and Whitehead B 1994 It's not easy being green *Harvard Business Review.* 36-44.

[4] Porter M and Linde Van 1995 Toward a new conception of the environment competitiveness relationship *The Journal of Economic Perspectives.***9** 97-118.

[5] Porter M 1996 America's green strategy *Business and the Environment* (London: Earthscan) p 33.

[6] Fu Jingyan and Li Lisha 2010 The empirical study of environmental regulation, factor endowments and industrial international competitiveness: based on panel data of China's manufacturing industry *Management Word*. **10** 87-98.

[7] Xiong Yan 2012 Environmental regulation on economic growth *Dongbei University of Finance and Economics* (Dalian).

TI 2000-100/3
Tinbergen Institute Discussion Paper

Ecological–Economic Analysis and Valuation of Biodiversity

Paulo A.L.D. Nunes
Jeroen C.J.M. van den Bergh
Peter Nijkamp

Tinbergen Institute
The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Keizersgracht 482
1017 EG Amsterdam
The Netherlands
Tel.: +31.(0)20.5513500
Fax: +31.(0)20.5513555

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31.(0)10.4088900
Fax: +31.(0)10.4089031

Most TI discussion papers can be downloaded at
http://www.tinbergen.nl

Ecological - Economic Analysis and Valuation of Biodiversity

Tinbergen Institute Discussion Paper

Paulo A. L. D. Nunes
Jeroen C. J. M. van den Bergh
Peter Nijkamp

Department of Spatial Economics
Free University
De Boelelaan 1105
1081 HV Amsterdam
Netherlands

e-mail: pnunes@econ.vu.nl, jbergh@econ.vu.nl, pnijkamp@econ.vu.nl

fax: +31.20.444.60.04

1. Introduction: context and scope of the study

In the spirit of the growing environmental awareness in the past decades, we witness nowadays an increasing interest in biodiversity, both locally and world-wide. Biodiversity requires our attention for two reasons. First, it provides a wide range of benefits to mankind and human activities. Second, many human activities have contributed to unprecedented rates of biodiversity loss, which threaten the stability and continuity of ecosystems as well as their provision of goods and services to mankind. Consequently, in recent years much attention has been directed towards the analysis and valuation of the (loss of) biodiversity. The valuation of biodiversity can be approached from an ecological, economic, or combined perspective. This contribution presents an overview of economic and ecological indicators of biodiversity, identifies the underlying valuation approaches, discusses basic concepts and theories, and reviews the respective applications.

The organisation of this study is as follows. Section 2 starts with a concise review of the most important frameworks that focus on the conceptualisation of biodiversity and its role in ecological processes. It continues then with the identification and discussion of alternative ecological valuation approaches of biodiversity, including biotic-richness and ecosystem health approaches. Section 3 addresses the economic concepts and methods relevant to the valuation of biodiversity. It starts with a discussion of the reasons why economists are interested in pursuing monetary value assessment of biodiversity and why some scientists disagree with that. A survey of empirical valuation studies is presented. The distinction between 'resource valuation' and 'biodiversity valuation' is discussed. Finally, respective valuation results and measurement methods are discussed. Section 4 identifies a combined economic-ecological valuation perspective. According to this perspective, biodiversity is analysed, modelled and valued from an integrated point of view, combining natural and social-economic aspects, and thus can enable the construction of mutually consistent indicators of biodiversity. Finally, Section 5 presents conclusions and outlines some recommendations for use of valuation methods and findings in environmental policy.

2. Ecological foundations for biodiversity analysis and valuation

2.1 Conceptualisation of biodiversity and its role in ecological processes

The analysis of biodiversity is rooted in the domain of both natural and social sciences and its modelling implies a review of knowledge on the relationships between biodiversity, the dynamics of ecosystems, and the level of human economic activities. One reason why biodiversity modelling has been so difficult relates to the complex and partly unobservable of the biodiversity-ecosystem relationships involved, such as biotic-abiotic interactions, food webs nutrient flows, and species interrelations. Independently of the complex nature of the biodiversity-ecosystem relationships, an important aspect is the recognition that the variability of the biological resources influences the functioning and structure of ecosystems. In the literature one can find three main approaches to modelling. These are reviewed in the present section.

Holling (1987 and 1992) proposed a model to describe and explain the dynamics of a terrestrial ecosystem in terms of a structure that is characterised by a sequential interaction between four basic functions or phases - the "4-box model". The functions are: (1) exploitation; (2) conservation; (3) release; and (4) reorganisation. Within this model,

ecosystems develop from the exploitation phase during which systems capture easily accessible resources, to the conservation phase during which systems build and store increasingly complex structure, and then evolve to the release phase during which systems free some of the mature structures. The released structure is then available for reorganisation and uptake in the phase of the exploitation phase. The exploitation function refers to the ecosystem's processes that are responsible for "colonising disturbed sites". The conservation function refers to the ecosystem's processes that are responsible for "resource accumulation that builds and stores energy and material". The release or "creative destruction" function refers to an abrupt change in the ecosystem caused by external disturbance, releasing energy and material that have been accumulated during the conservation phase. Examples of the release phase are fire, storms, and pests (Costanza *et al.* 1995). Finally, the reorganisation function refers to the ecosystem's processes that are responsible for mobilising released energy and materials and making them available for the next exploitative phase.

Based on a categorisation of the ecosystem's functions initially made by Odum (1971), de Groot (1994) describes the relationship between biodiversity and ecosystem. In general terms, de Groot characterises the ecosystem structure in terms of four categories of biodiversity functions: (1) life support functions, (2) carrier functions, (3) production functions, and (4) information functions. Biodiversity is seen to have a life support function, i.e., a regulation of essential ecological processes. The life support functions refer to the group of the biodiversity service flows that contribute to the maintenance of a healthy environment, by providing clean air, water and soil, by providing flood control, and by providing carbon storage and waste absorption. Most of the life support functions are often fuzzy (e.g., provision of carbon storage), and thus not easily to be demarcated and identified. The carrier functions refer to the provision of space for human activities such as habitation, agriculture, and recreational activities. The production functions refer to the provision of environmental resources, ranging from raw materials for industrial use to water and energy resources. The information functions refer to the maintenance of mental health, providing opportunities for reflection, spiritual enrichment and aesthetic experience.

More recently, Norberg (1999) proposed an alternative approach to classify ecosystem functions and services of Nature. Norberg selected groups of ecosystem services to which common ecological concepts apply: (1) are the goods and the services internal to the ecosystem or shared with other systems?, or, (2) are the goods and the services of biotic or abiotic origin?, and (3) at which level of ecological hierarchy are goods and services maintained? Bearing in mind such selection criteria, ecosystem functions and services of Nature are classified in three categories: (1) maintenance of the populations; (2) regulation of material and energy flows; and (3) organisation of biological units through selective processes. These categories represent three major fields in ecology that have well-established theoretical foundations and refer to population/community ecology research, ecosystem research, and biological organisation, respectively (Levin *et al.* 1997, Levin 1998).

The first category corresponds to the group of ecosystem services that are "(…) associated with certain species or a group of similar species (…)" (Norberg, page 185). Examples of such services include valuable foods and goods as fish, timber, pharmaceuticals chemicals, and flowers. The second category consists of processes that regulate exogenous chemical or physical cycles, i.e. processes that drive material and energy flows in ecosystems. The biota takes a significant part in most global cycles of chemical compounds such as water, CO_2, and nitrogen. Finally, the third category of ecosystem services is related to the organisation of

biotic entities. Organisation is virtually present at all scales: organisation of genes through natural selection, spatial distribution of a population through dispersal and competitive exclusion, or the development of food webs and ecosystems through invasion and extinction processes.

2.2 Ecological indicators of biodiversity: (1) biotic-richness approach

From a biotic approach, the definition of biodiversity 'value' indicates the range magnitude of biological products and services flows provided by nature. Traditionally, the measurement of biological diversity has been undertaken with the use of genetic, species, and ecosystem richness or variety indices.

Measurement of genetic diversity

The analysis, conceptualisation and measurement of genetic differences can be done in terms of (1) allelic frequencies, (2) phenotypic traits and (3) DNA sequences. The same gene can exist in different frequencies or variants. These variants are called alleles. Thus, allelic diversity measures the different gene composition variants of individuals. In general, the more alleles, the more diverse their frequencies, the greater the genetic diversity. Average expected heterozygosity, the probability that two alleles sampled at random are different, is commonly used as an overall measure. A number of different indices can be applied to the measurements to assess average expected heterozygosity (Antonovic 1990).

Phenetic diversity is a measure based on individual's phenotype traits, checking whether individuals share the same characteristics. This valuation method is focused on the measurement of the variance of certain traits and, in general, involves readily measurable morphological and physiological characteristics of the individual. However, individual genetic information is often difficult to assess, and comparisons are difficult, when the individuals or populations are measured in terms of different qualitative traits. To cope with that, scientists may now use DNA sequence variation so as to measure the genetic variety. The DNA sequence information is obtained by the use of the polymerase chain reaction. For this reason, only a small amount of material, ultimately one single cell, is required to obtain the DNA sequence data. Closely related species may share up to 95% of their DNA sequences, thus implying little diversity in the overall genetic information.

Measurement of species diversity

The measurement of species diversity, in its ideal form, consists of a complete catalogue of the distribution and abundance of all species in the area under consideration. However, this measurement is usually not possible unless the area under consideration is a small area. Therefore, in practice the measurement of species diversity is often based on samples. The central measures of species diversity are the α, β and γ species diversity, as originally proposed by Whittaker (1960 and 1972). α-diversity refers to the number of species using only their presence (and not abundance) in a given area. It therefore measures the species richness of a given sample plot. The use such a species diversity measure would imply that an area containing a high number of species would be preferred to one with a fractionally smaller number of species (Huston 1994).

Another measure of species richness is γ-diversity. It is often used to assess the overall diversity within a large region and its comprehension has direct implications when dealing with biodiversity at the landscape level (Noss 1983, Franklin 1993). National species lists,

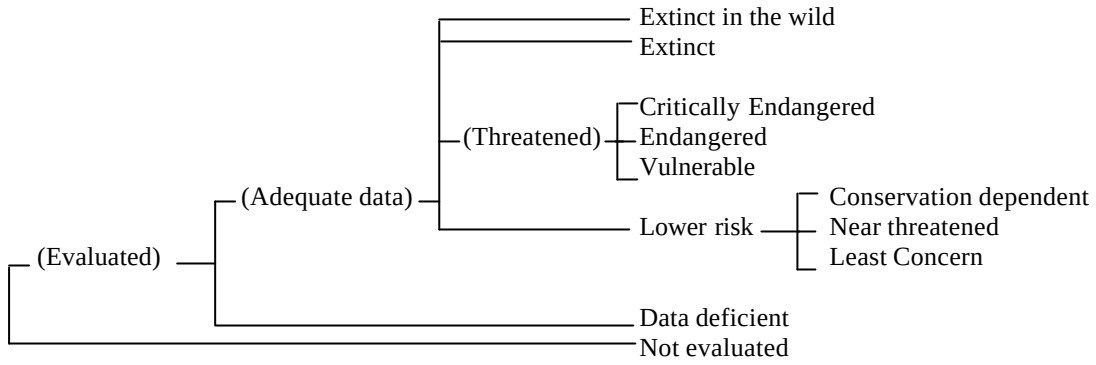
usually the only information available, can be treated as lower bounds on gamma diversity. Colombia and Kenya, for example, are the homes to over 1,000 species of birds, while the UK and the forests of eastern North America are homes to only about 200. A coral reef of northern Australia may be a home to 500 species, while the rocky shoreline of Japan may be a home to only 100 species (UNEP 1995). Finally, b-diversity measures the turnover of species between local areas, i.e. the rate of change in species composition among discrete sites or habitat units (Cody 1986 and 1993). As such it cannot be expressed in the number of species: it is represented in terms of an index and it is interpreted as a species turnover rate. b-diversity is generally used to estimate average changes in species in response to site or habitat heterogeneity.

Species richness measurement is useful, but it may depict a biased diversity assessment. First, scientists face great uncertainties about the total number of species. Therefore, species richness can be only be measured for some species (the species studied), but not for all the species in the area. As a matter of fact, only in a very few places on the planet are there rough estimates for the total number of species. Second, the size of the area is often arbitrary. Species diversity is associated to habitat scale in a complex way. Thus, one needs to be cautious when comparing the species diversity of areas that differ greatly in size. In addition, species diversity is a result of complex genealogical relationships that are not measured here. Alternative species diversity measures supplement species richness with measures of the degree of genealogical difference. Such diversity measures include the weighting of close-to-root species, higher-taxon richness, spanning tree length and taxomonic dispersion (Williams *et al.* 1991). However, and for the time being, practical difficulties with respect to the implementation of such measures confer reliance on the simpler indicators of species richness.

Measurement of ecosystem diversity

The measurement of biodiversity at the organisational level of ecosystem diversity encompasses a multi-complexity of relationships, both at the intra- and supra-species level, that play a crucial role in defining the overall distribution and abundance of species. For this reason, there are a number of factors that make the assessment of ecosystem diversity more blur and less clearly defined. Actually, at the ecosystem level many different units of diversity are involved, ranging from the patterns of habitats to the age structure of populations, including the patterns of communities on the landscape and patch dynamics. At these levels it is not clear where to draw boundaries delineating units of biodiversity. When a wetland, for example, is disturbed or degraded, we need to look at the impacts of the disturbance on the larger level of the landscape. Emphasis on a system-wide approach also serves to remind that the ecological value of an ecosystem may not be equivalent to the aggregate value of that same system's components. In other words, the system is more than just the aggregation of its individual parts; it possesses a primary value (Gren *et al.* 1994; Turner, Perrings and Folke, 1997). Furthermore, the conservation of biodiversity at the ecosystem level does not only underlie the preservation of species, but also contributes to the safeguard of the ecosystem functions and services. Thus, the full range of biodiversity values depends on the processes that support the functioning of larger-scale ecological systems. Given such unambiguous boundaries, there are different measurement approaches. These include bio-geographical realms or provinces, based on the distribution of species, and eco-regions or eco-zones, based on physical attributes such as soils and climates (UNEP 1995).

Table 1: Structure of species categories



Source: http://www.iucn.org/themes/ssc/redlists/categor.htm

Operationalisation of the biotic-richness approach: examples

Ecologists are often asked to contribute with their expertise and help policy makers to define conservation priorities. A considerable part of the ecologist's help is, in some way, related to Usher's ecological approach to environmental protection (Usher 1989). According to Usher's conservation assessment, decisions are characterised by three steps. First, attributes are identified and are used to reflect the conservation interest of the species or site. Second, criteria are developed for the expression of the attributes in a form that allows evaluation. Finally, values are attached to particular levels of criteria.

One important instrument for assessing species variety is the Red Data Book (e.g. ICN 1993). Since their introduction, Red Data Books were conceived as to identify threats or causes of decline of different species around the world (Fitter and Fitter 1987, Mace and Lande 1991, Mace *et al.* 1992, IUCN 1993, Mace and Stuart 1994). In short, Red Data Book is an evaluation technique characterised by the selection of a species list (the attribute) to assess the species richness (the criterion). Red Data Books classify species in one of eight different categories: extinct, extinct in wild, critically endangered, endangered, vulnerable, lower risk, data deficient and not evaluated – see Table 1. The goal is to provide an easily and widely understood method to order species in categories related to their threat of extinction under current circumstances[1] in order to (1) provide information on which to base conservation programs; (2) assist the drafting of legislation, and (3) to convey information comprehensible to a non-specialist. For this reason, Red Data Books are frequently used by numerous governmental and non-governmental organisations for policy guidance and the establishment of conservation priorities.[2]

Several problems however exist, rendering category assessment quite difficult to operationalise. While, the classification of a species to a category relies on an objective

[1] For example, "critically endangered" category refers to the species that is facing an extremely high risk of extinction in the wild in the immediate future. Alternatively, a species is "endangered" when it is facing a very high risk of extinction in the wild in the near future. Finally, a species is said to be "vulnerable" when it is not "critically endangered" nor "endangered" but yet is facing a high risk of extinction in the wild in the medium-term future (see Annex for more details).

[2] Most obviously these were required for selecting and designating the European Union 8,819 nature reserves under the Natura 2000 network, 6,977 within the Habitat Directive (92/43/CEE), and 1,842 within the Birds Directive (79/409/CEE).

evaluation, the actual definitions of these categories rely on a subjective view. In practice, the large number of available criteria (e.g. a , b and g criterion) used for evaluation in some way already reflect the difficulties that exist in conceptualising its value. Furthermore, the species category of threat is not necessarily sufficient to determine priorities for conservation action. Therefore, if Red Data Books were elected as the ecological method used in the OECD countries for the establishment of biodiversity priorities, it would be a risk that many sites would not have much of a value (and thus more likely not to be protected), since this evaluation technique simply provides an assessment of the likelihood of species extinction. Finally, given the scientific understanding of population and ecosystems, it is possible to develop alternative indicators, including the use of numerous other criteria concerning conservation action.

An early example of a multicriteria rating is the method proposed by Randwell (1969). The method was used to evaluate coastal habitats and combining the use of eight criteria into a single score, the Comparative Biological Value Index (CBVI). Each of these criteria are rated according the scale as described in Table 2 and the final score is obtained by summing up the scores for all the nine criteria:

$$CBVI = Ph + O + D + G + S + P + E + C$$

The maximum potential value is 28 and the minimum value is 7. The higher the CBVI value, the greater is the requirement for site protection. Since Randwell, the use of indices constitutes a popular practice in ecological valuation and management (see Spellerberg 1992 for an extensive review of CBVI assessment of landscape and urban habitats). However, this valuation approach relies on input criteria that require some subjective valuation. As such they may not be helpful to the decision-making process as originally intended. This problem is confounded as assessing and representing ecological values for policy guidance.

Recently, however, computer-based systems have been used to develop a general, integrated framework where one can simulate natural/management changes and assess the respective conservation evaluations. One specific output of this evaluation approach is to study population dynamics and estimate a minimum viable population, defined as the smallest population which as an acceptable probability of persisting over a given time period (Soulé 1987).[3] This in turn allows the calculation of the minimum dynamic area, i.e. the geographic area of suitable habitat required as to support the minimum viable population. Recent published studies on the leadbeaster's possum (Lindenmayer *et al.* 1993), the eastern barred bandicoot (Lacy and Clark 1990), the Puerto Rico parrot (Howels and Edward-Jones 1997), and the giant panda (Zhou and Pan 1997) confirm the use of the population viability analysis in the policy decision circles and help conservation decisions. Another example of the application of computer system for ecological evaluation refers to the System for Evaluating Rivers for Conservation (SERCON) recently undertaken by the Scottish Natural Heritage, the nature conservation agency in Scotland (Boon *et al.* 1997). The overall objective was to predict the impact of different development scenarios on river ecological conservation value as well as to provide of a simple way of communicating such results to planners, developers and policy-makers.

[3] Several general models have been developed for this task, e.g. VORTEX (Lacy 1993) and METAPOP (Akcakaya 1994), and these can be used to consider the complex interactions between demographic, environmental and genetic influences on a population.

Table 2: Rating of the criteria used by Randwell for evaluating coastal habitats

Criteria	Type		Rating
Pysicochemical featuares (Ph)	High speciality		3
	Some special features		2
	Type example		1
Optimum populations (O)	Best populations of one or more local species		4
	Large populations of local species		3
	Large populations of common species and small populations of local species		2
Diversity (D)	Representative populations		1
	Outstanding populations		3
	High diversity		2
	Species ranges small		1
Geographic units (G)	Many species at limit		3
	Some species at limit		2
	Dew species or no species at limit		1
Size (S)	Mud-flats (ha)	Cliffs (km)	
	> 4000	> 80	5
	1600 – 3999	40-79	4
	800 – 1599	24-39	3
	400 – 799	8-23	2
	< 400	<8	1
Purity (P)	Little disturbance		3
	Moderate disturbance		2
	Much ground disturbed or polluted		1
Education and research use (E)	Much used		3
	Some use		2
	Potential use		1
Combinatory value (C)	Adjacent to another habitat of likely national value		4
	Adjacent to another habitat of likely regional value		3
	Adjacent to another coast habitat site not spoilt by development		1
	Surrounded by developed coastline		0

Source: Randwell (1969)

Ecological valuation based on computer modelling presents important advantages. First, it has encouraged greater rigour in data evaluation, since it permits the introduction of elements of subjectivity within a transparent and repeatable framework. Second, it allows a direct comparison of alternative conservation policies, independently of the number of the involved criteria and respective attributes (e.g. it permits the comparison of a conservation policy involving a criterion with ten attributes which could score a maximum of 50, with another strategy involving a criterion with two attributes that could only score a maximum of 10). Finally, it permits an evaluation even when some of the attribute data are missing, which is quite common in practice.

2.3 Ecological indicators of biodiversity: (2) ecosystem health/integrity approach

Ecological valuation methods are not only aimed at assessing diversity and rarity of species, but also the complex interactions between the biotic and abiotic environments, based on the assumption that the variety of abiotic conditions is equally important as variety of species. For instance, abiotic diversity (e.g. physical characteristics of the landscape such as soil pH

and salinity) is expected to be linked to the prevalence of endemic species and thus to biotic diversity and rarity in a natural way (Bertollo 1998). Therefore, from an ecosystem perspective, the definition of biodiversity value is intrinsically related to ecosystem performance and integrity. Furthermore, the term "value" indicates how well an ecosystem is functioning when compared to its own potential and how important this is for the functioning of other ecosystems and, ultimately, for the functioning of the global ecosystem (Sijtsma *et al.* 1998).[4]

Ecosystem health

Ecosystem health is an overall indicator of the ecosystem functioning (or ecosystem integrity), taking into account both ecological and human processes. An ecological system is said to be healthy, if it is stable and sustainable, i.e. if it is active and maintains its organisation and vigour over time and is resilient to stress. In short then, ecosystem health can be defined as "(…) a measure of the overall performance of a complex system that is built up from the behaviour of its parts (…)" (Costanza 1992, pp. 241). Before we can measure the health of an ecosystem it is, however, necessary to proceed to identify biotic and abiotic parameters or indicators, target human economic activities, and the scale or hierarchy of analysis - see Figure 1.

When selecting the parameters, and indicators, it is important that these are relevant to the analysis. Most of the biotic and abiotic parameters or indicators, such as indexes describing the soil, flora, and fauna, have emerged from the ecological literature (Odum 1971). Furthermore, the measurement of the selected biotic and abiotic indicators needs to be a feasible task, so that it is possible to come up with plausible, valid figures. In addition, it is necessary to frame the scale, or hierarchy, of analysis. The choice of the scale relates to important decisions over boundary and temporal perspective of analysis (Norton and Ulanowicz 1992). It is common that boundaries are drawn in accordance with the ecosystem's land features or its geography (e.g. wetlands ecosystem). Finally, the measurement of ecosystem health also requires the identification of the human actions that influence the ecological structure and processes. The underlining idea is that human economic activity or target groups, e.g. consumers, industry, influence their environment. As a result, information on the impact of economic activity on the ecological structure and processes, in general, and on economic indicators, in particular, needs to be taken into account when assessing the ecosystem health.

It is also clear that both ecological and human dimensions are dynamic; they change in different ways and according to multiple frames of time. Consequently, biotic, abiotic, and economic indicators must be sufficiently dynamic to change accordingly. For example, even if we have clearly defined the spatial and temporal boundaries of analysis, there could remain important questions regarding the structure of the ecosystem, originally bounded and described as a whole. Therefore, there would be the need to represent boundaries at a smaller scale. Each part is characterised by its own set of indicators and is assessed individually. This set of indicators can differ substantially from part to part, i.e., from ecosystem to ecosystem (Costanza *et al.* 1992).

[4] From an ecological perspective, the definition of 'value' indicates how well an ecosystem is functioning compared to its own potential (Sijtsma *et al.* 1998). Furthermore, if two ecosystems functioning equally well with respect to their own potential, one system may be more valuable because of its positive synergies to the well-functioning of other systems.

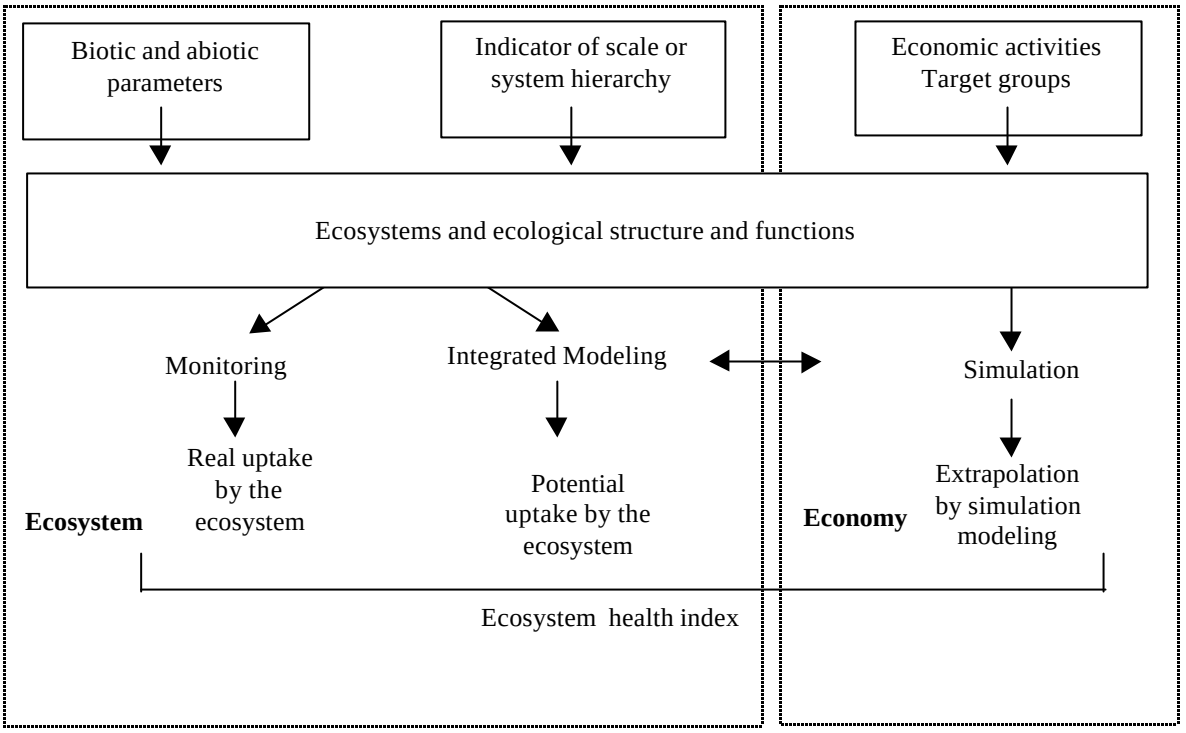


Figure 1: Definition of ecosystem health indexes

After having decided upon the biotic and abiotic indicators, scale or hierarchy of analysis and the target economic groups, the scientist is in a good position to proceed with the measurement of the overall ecosystem performance, i.e. ecosystem health. One possible approach is to use directly the available data provided by monitoring activities. This data will provide information with respect to the real uptake by the ecosystem and this way allow to measure the overall ecosystem health. Alternatively, one can integrate the relevant concepts and develop an analytical framework. The combination of the available data and such a model formulation will allow for a further step in the assessment of ecosystem health. At this stage, the scientist is not only to use the available data to deduce ecosystem health, but also explore the dynamics of integrated modelling so as to estimate the potential uptake by the ecosystem. Finally, the scientist is able to simulate different conservation scenarios by manipulating the observed characteristics, or 'controlling variables' - e.g., introduction of a new set of concentration ratios. For each conservation scenario the scientist is able to compute the associated ecosystem health index and thus provide crucial information as to rank the alternative management policy scenarios.

Examples of ecosystem health indicators

The construction of ecosystem health indexes allows policy makers to predict ecosystem response as a result of various specific management alternatives and natural changes. In practice, this formulation leads to the following form for an overall system health index (HI):

$$HI = V*O*R$$

Table 3: Indices of vigour, organisation and resilience

Components of health	Related concepts	Related measures	Field of origin	Measurement solution
Vigour	Function	GPP, NPP, GEP	Ecology	Monitoring
	Productivity	GNP	Economics	
	Throughput	Metabolism	Ecology	
Organisation	Structure	Diversity index	Ecology	Network analysis
	Biodiversity	Mutual information predictability	Ecology	
Resilience		Scope for growth	Ecology	Simulation modeling

Source: Costanza (1922)

where *V* designates system 'Vigour' and represents a cardinal measure of the ecological system activity, metabolism, or primary productivity; *O* designates system 'Organisation' and represents a 0-1 index of the relative degree of the ecological system's organisation, including its diversity and connectivity; and *R* designates system 'Resilience' and represents a 0-1 index of the relative degree of the ecological system's resilience – see the components in Table 3. In short, and to conclude, the overall ecological system health is given by its activity weight through indices for the relative system's organisation and resilience. To operationalise the vigour, organisation and resilience components, the health index will require the application of different measurement solutions to current data, involving the use of expertise from both economics and ecology.

Ulanowicz's ascendency index

One important example of ecosystem-health index, Ulanowicz's ascendency index, allows for an integrated, quantitative, hierarchical measurement of ecosystem health (Ulanowicz 1992). In simple terms, the ascendency index reflects any degradation of the system. However, such an indicator requires data on all transfers occurring in the ecosystem under consideration. The collection of this data is usually a laborious, expensive task. For this reason, fully quantified networks of ecosystems still remain scarce (Costanza 1992). In addition, it is important to remember the distinction between a scientific-orientated methodology and a policy-orientated approach. In practice one can find several valuation methods that combine a set of ecosystem integrity indicators into one common denominator that has a socio-politically appeal and thus a meaning for the policy makers. The following sub-sections review the operationalisation of some of the ecological valuation methods, most of them with an application to the Dutch regional and national policy agenda (see Ruijgrok 1999).[5]

The ecosystem classification method

Brink and Hosper (1989) developed a General Method for the Description and Evaluation of Ecosystems, known by the Dutch acronym AMOEBE.[6] The method was originally used to assess the quality of aquatic ecosystems by comparing the presence of selected species with a their presence in a benchmark situation of 1930. The selected species, which Brink and

[5] The characteristics of the Netherlands, a country with the highest population density in the European Union and home of important human economic activity, urge the use of this ecological valuation methods more than in any country.

[6] Algemene Methode voor OEcosysteembeschrijving en Beooordeling.

Hosper called ‘target variables’, were selected on the basis of: (1) their representativeness (i.e. do they represent a healthy aquatic ecosystem); (2) their flexibility (i.e. can they be influenced by human interventions), and, finally (3) their measurability and data availability (i.e. are they easy to measure and/or are there data-bases available). Since the AMOEBE does not indicate whether one ecosystem is more valuable than another (the value of the ecosystem is a function as a deviation from its own potential), the valuation method is not often selected with the goal of policy formulation guidance.

Nature measurement method

In 1995, the Dutch Centre for Agriculture and Environment in Utrecht developed the nature measurement method as to assess the natural values of agricultural areas. The natural values were measured in terms of species abundance and its deviation from its own diversity potential. Species were selected that are present on agricultural plots, easy to recognise and that represent natural quality (Buys 1995). A similar formula was developed by the Foundation for Spatial Economics of the University of Groningen to assess the costs and benefits of the National Ecological Network (Sijtsma and Strijker 1995). The costs were valued in monetary terms and the benefits mostly in ecological terms. The National Ecological Network used the identification of ‘nature target types’ (i.e. pre-defined types of nature such as the European CORINE network) that are assumed to provide the habitat mosaic for ‘target species’. To this purpose, digital thematic maps are generated as geographic information systems land cover habitats types. Target species are, in turn, classified and selected on basis of national and international rarity. Zurlini, Amadio and Rossi have recently applied this valuation method to depict the map of Italian nature (see Zurlini *et al.* 2000).

Ecological effect measurement method

In 1996, the Centre for Environmental Studies in Leiden developed an ecological effect measurement method to value the effects of housing development projects on nature and landscape. The method is as a first step characterised by the description of the reference situation and human intervention measures, followed by the determination of the effects on nature and, finally, proceeding to the aggregation of effects (Cuperus and Canters 1995). The main ecosystem biotic and abiotic characteristics were used for valuation, including spatial diversity (e.g. variables on soil and vegetation structures), abiotic functioning (e.g. variables on temperature, sediments, water and soil), fauna and flora communities as well as their relations with the surroundings (e.g. variables on hydrologic, geomorphologic and biomass relations). Since this valuation method takes into account both biotic and abiotic diversity assessed by means of the deviation from a reference situation, it can easily be transferred to the policy arena and used, for example, to determine compensation measures in the case of damage to existing natural areas.

Ecological capital index

More recently, the Dutch Environmental Planning Bureau in Bilthoven developed the ecological capital index with the objective to assess the state of both natural and cultural ecosystems in relation to human activities. This ecological capital index is calculated by multiplying the ecosystem’s quantity by its quality The abiotic environment is here regarded as a conditional variable for such a biodiversity reference situation (RIZA 1999). The

international application of the ecological capital index respects the recommendations for a core set of indicators as proposed by the Convention on Biological Diversity (UNEP 1997) and thus is compatible with the international classification (IUCN 1991) of ecosystems on the basis of the degree of human influence.[7]

To conclude, ecosystem health indexes allow the scientist to assess the overall ecosystem performance. A number of examples have been identified here. These play a crucial for policy guidance since they let us to predict ecosystem response as a result of alternative management scenarios and, this way, making possible scenario's rankings to be compared.

3. Economic foundations for biodiversity analysis and valuation

3.1 Why economists pursue economic valuation

The economic valuation of natural resources, in general, and biodiversity, in particular, is among the most pressing and challenging issues confronting today's environmental economists. In short, economists value biodiversity because such a valuation exercise allows for a direct comparison with economic values of alternative options and facilitates, for example, cost-benefit analysis – a crucial tool for policy formulation. In addition, the monetary valuation of biodiversity allows economists to perform environmental accounting, natural resource damage assessment, and to carry out proper pricing. Moreover, the valuation exercise is revealed to be essential in the research of individual consumer behaviour and investigate what does the individual consumer think of certain biodiversity management objectives, or identify individual consumer's motivations with respect to biodiversity conservation.

Many people, however, do not accept to a place monetary values on biodiversity. Arguments against it are rooted in the human preference orientation that 'guide' consumer behaviour with respect to biodiversity (Ehrenfeld 1988, Lockwood 1999). At the risk of oversimplification, we can distinguish two broad ranges of value orientations - see Table 4. According to the 'anthropocentric' orientation, the value of biodiversity is an outcome of its role in human welfare, as humans conceive it "(…) whether they are selfish, altruistic, loyal, spiteful, or masochistic (…)" (Becker 1993). A second valuation perspective is rooted in an 'biocentric' or 'ecocentric' value orientation, which claims that nature has an intrinsic value[8] and therefore deserves protection. In reality, however, value orientations are overlapping and several versions of 'anthropocentrism' and 'ecocentrism' can exist within one individual. Q-altruism and stewardship are examples of such 'mixed' attitudes (Sagoff 1980, Norton 1982, Van der Veer and Pearce 1986). Stewardship is a form of altruism that is fully divorced from any explicit notion of consumption. It corresponds to a sense of responsibility – usually in a Christian perspective - for the conservation and maintenance of the resource. Q-altruism is rooted in the firm belief that living organisms are incapable of protecting themselves against human actions. Therefore, the conservation of living organisms merits humans' sympathy or compassion.[9]

---

[7] The International Union for Nature Conservation, under the auspices of the United Nations, makes a distinction between natural, adapted, cultivated, built and deteriorated ecosystem.

[8] By definition an intrinsic value is the value of the resource *per se* without a subject attaching a value to it.

[9] Recently, Nunes (1999a) developed the use of multiple motivation items scales to measure consumers' level of "anthropocentricism" and integrate such a construct in a contingent valuation application. The empirical results confirmed that the individual level of "anthropocentricism" had a robust econometric role in the economic

Table 4: Value orientations and environmental attitudes

Value orientation	Valuation perspective	Ethical approach	Biodiversity attitude
Ecocentrism	rights conferred to all living organisms	Nature has intrinsic value, regardless of human recognition	Biodiversity first
Antrophocentrism	rights and interests conferred to individual humans	Value of nature is value conferred to humans	Humans first

Adapted from Lockwood (1999)

3.2 Biodiversity as a source of economic values

The concept of total economic value of an environmental resource has its foundations in welfare economics. It focuses on changes in the economic welfare of humans. Therefore, the terms 'economic value' and 'welfare change' can be used interchangeably.[10] The total economic value (TEV) of an environmental resource consists of its use value (UV) and nonuse value (NUV). Use values arise from the actual use. It can be further divided into direct use values (DUV), indirect use values (IUV) and option values (OV) (Pearce and Moran 1994). The nonuse values are usually divided into a bequest value (BV) and an existence value (XV). Bequest value refers to the benefit accruing to any individual from the knowledge that others might benefit from an environmental resource in the future. Existence value refers to the benefit derived simply from the knowledge that the resource is protected. This leads to the following equation

$$TEV = UV + NUV = (DUV + IUV + OV) + (XV + BV)$$

Economic values of changes in biodiversity require special attention. We classify such values into four categories. These are illustrated in Figure 2. A first category, denoted by link 1 6, depicts biodiversity benefits in terms of ecosystem functions and services (see Section 2.1). We can interpret this biodiversity value category as a direct use value component. A second category, denoted by link 1 4 5, captures the value of biodiversity in terms of supply of ecosystem space or natural habitat protection. We can interpret this biodiversity value category as an indirect use value component. A third category, denoted by link 2 5, captures the benefits conferred to society in terms of an overall diversity provision of biological resources, notably specific animal and species for use in agriculture and medicine. We can interpret this biodiversity value category as a direct use value component. A fourth category, captured by link 3, refers to the direct impact of biodiversity on human welfare. The economic value of biodiversity is then measured in terms of philanthropic considerations, independently of biodiversity use or consumption. One can interpret this biodiversity value category as a nonuse component. The nonuse values have a public good character for which no market price is available. As a consequence, most of the time policy makers have based their decisions on an undervaluation of biodiversity benefits, which has resulted in a misuse

valuation exercise. Moreover, the respective coefficient estimate show to vary with the object submitted to valuation.

[10] It should be clear that economists do not pursue total value assessment of an environmental system, but of system changes (preferably marginal or small).

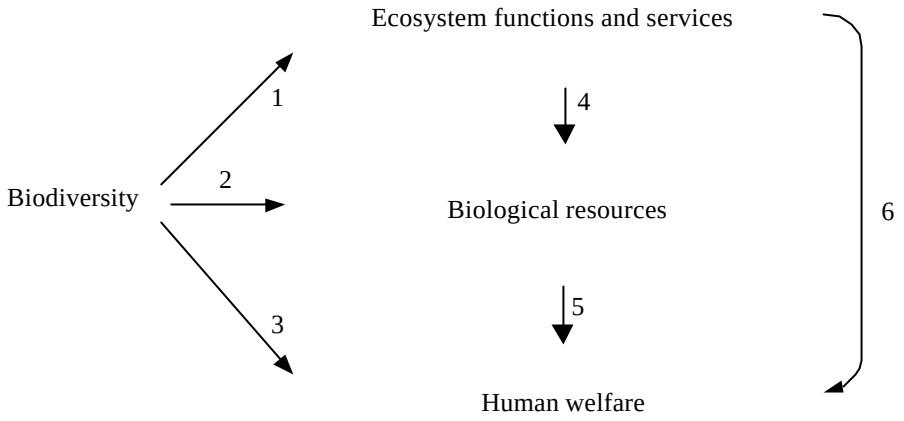


Figure 2: Economic values of biodiversity

and misallocation of scarce resources. The monetary assessment of biodiversity use and nonuse benefits requires the handling of special valuation tools. These are discussed in the following section.

3.3 Alternative monetary valuation methods

In the absence of market prices, certain techniques are needed to retrieve consumers' preferences for biodiversity benefits. On the basis of the biodiversity component to be valued, one can distinguish alternative valuation methods: travel cost, hedonic price, averting expenditures cost, production function, and contingent valuation,[11] - see Table 5. Travel cost is also a frequently used economic valuation method (Bockstael *et al.* 1991). It is especially designed to assess recreational benefits by looking at the generalised travel costs (site entrance fees, accommodation and journey costs) when visiting a natural park or reserves. Alternatively, when using the hedonic price method (Palmquist 1991) to estimate the economic value of a biodiversity benefit, say, good soil conditions, researchers explore the analysis of land market values and the soil quality characteristics (e.g. depth of topsoil). For example, King and Sinden (1988) explored the hedonic price method in order to capture the value of soil conservation in the farm land market of Manilla Shire, in New South Wales, Australia. The hedonic land market price regression results show that the implicit marginal price of land condition is estimated to be $2.28/ha. Researchers can also estimate the economic value of biodiversity on the basis of consumer's expenditures made to avert or mitigate the adverse effects (Cropper and Freeman 1991) derived from the loss of biodiversity benefits. Huszar (1989) studied wind erosion costs that New Mexico households' incur from increased cleaning, maintenance and replacement expenditures, and reduced consumption and production opportunities. In addition, the economic value of biodiversity can be assessed through a production function. Recently, Barbier (1994) conducted a study focused on the value assessment of the economic importance of Hadejia-Jama'are wetlands, Nigeria. The valuation is based on the estimation of some key direct use values that the floodplain provides to local population through crop production, fuelwood and fishing.

[11] Many other classifications of valuation exist. See Freeman (1979).

Table 5: Identification of monetary valuation methods according to the different biodiversity value components

TOTAL ECONOMIC VALUE (TEV)	USE VALUES (UV)	Direct use value (DUV)	Recreation benefits, e.g. sight-seeing, fishing, swimming Methods: Travel cost, contingent valuation
		Indirect use value (IUV)	Ecosystem functional benefits, e.g. regulating local chemical composition of the water Methods: production function, averting behaviour, hedonic price
		Option Value (OV)	Insurance for having the asset on *stand-by*, e.g. future visits, future genetic manipulation Methods: contingent valuation
	NONUSE VALUES (NUV)	Bequest Value (BV)	Legacy benefits e.g. habitat conservation for future generations Methods: contingent valuation
		Existence value (XV)	Existence benefits, e.g. knowledge of existence of marine wildlife diversity Methods: travel cost, contingent valuation

Finally, the contingent valuation method collects data to retrieve individuals' preferences on biodiversity by means of questionnaires, nowadays often supported by multidisciplinary research teams (Carson *et al.* 1992, NOAA 1993). Contingent valuation presents important features. First, it is able to express the Hicksian welfare measure directly in monetary terms. Second, the CV method is the only valuation technique which is capable of shedding light on the monetary valuation of the nonuse values, i.e., the benefit value component that is not directly associated with its direct use or consumption. Ignoring such values will be responsible for a systematic bias in the estimation (essentially an underestimation) of the total benefits of biodiversity. Moreover, CV brings with it the advantage that environmental changes may be valued, even if they have not yet occurred (*ex ante* valuation), offering a greater potential scope and flexibility in specifying different states of nature that may even lie outside the current institutional arrangements or levels of provision. For the same reasons, CV can be used to assess almost all value components of biodiversity and thus is frequently used in environmental policy.[12] Contingent, and other, valuation studies will be reviewed and discussed in more detail in the following section.

3.4 Empirical valuation studies

This section reviews some biodiversity valuation studies, presenting a valuation range for each biodiversity value component.[13] The discussion is organised according to Figure 2. First, we focus on studies that perform a value assessment of biodiversity in terms of the benefits associated with the provision of species diversity (see link 2 5). Second, we focus on valuation studies that pursue the assessment of biodiversity benefits in terms of protection of natural habitat or ecosystem space (see link 1 4 5). Third, we discuss valuation studies that focus on biodiversity's benefits conferred to the society in terms of ecosystem functions and

[12] An exception is made for the ecosystem functional benefits since these are hard to describe in a questionnaire format and they are often related to issues not particularly familiar to most respondents.

services, see link 1 6. Finally, whenever the CV method is applied, one is able to assess the nonuse value component of biodiversity, and thus capture the monetary value of link 3.[14]

Single, multiple species valuation surveys

One of the first species valuation studies was conducted by Boyle and Bishop in 1987. They conducted a CV study to estimate the existence values for wildlife species. The estimated values suggest that substantial existence values are associated with the Bighorn Sheep in Wyoming (see results in Table 6). More recently, Van Kooten (1993) studied the economic value of waterfowl in Canada. According to the results, the shadow values of marginal land converted to waterfowl habitat were estimated to be $50 to $60 per acre. Loomis and Larson (1994) valued an "emblematic" endangered species, namely the Gray Whale. The willingness to pay (WTP) was estimated between sixteen and eighteen dollars per household. Boman and Bosdedt (1995) carried out the economic valuation of the conservation of the Wolf in Sweden. The estimation results show that the mean WTP is about $126 per year.

Table 6 shows also multiple species valuation studies. Johnansson (1989) conducted a CV study that focus on the preservation of 300 endangered species in Sweden. As we can see, the final estimate, $ 194 per year, is higher than economic valuation of the conservation of the Wolf in Sweden, about $126 per year, though not so high as one would expect. This is partly because the single species valuation study refers to popular, socially attractive, charming and beautiful species. Therefore, when preserving the Wolf *in situ*, as one of the faunal emblems of Sweden, the final value estimate may also embed other characteristics that not directly related to the species to be valued. In other words, the valuation results may also reflect individual social esteem motives or warm glow (Nunes 2000a).[15] Furthermore, the estimate valuation for single species can be affected by the availability of related species, i.e. single species values can be affected by the respondent's perception of substitutes and complementary species (Samples and Hollyer 1989). For example, preserving a Wolf *in situ* also preserves the habitat for other endangered species; namely the ones that share the same environment as a Wolf. Therefore, the WTP for Wolf may be interpreted as a result embedding a valuation of a wider range of species. For this reason, some authors prefer to work with a higher aggregation level of species diversity and carry out valuation surveys at a habitat protection level.

Diversity at the ecosystem level

The majority of CV studies that focus on of biodiversity at the ecosystem level link directly the valuation of biodiversity to the non-use or recreational valuation of habitat protection programs. The main reason for such a link is primarily the difficulty associated with defining in a survey such an abstract concept such as ecosystem diversity. Indeed, some CV studies indicate that the concept of biodiversity is ill understood among the general population (Hanley *et al.* 1995). However, a number of valuation studies have attempted to value biodiversity conservation policies by using other methods. Generally speaking, we can find

[13] Nunes and van den Bergh (2000) list a more extensive bibliography of work related to valuation of biodiversity.

[14] Since nonuse values have no behavioural market trace, economists cannot retrieve information about these values by relying on market-based valuation approaches.

[15] This is a well-known critique supported by economists such as Jerry Hausman, Peter Diamond, William Desvousges and Paul Milgrom, who express their doubts with respect to the suitability of CV results for inclusion in benefit-cost analyses (Hausman 1993).

Table 6: Valuation studies

	Valuation study	Method
Single species		
Minimum range: $5	Stripped Shiner, endangered species in Wisconsin, US Boyle and Bishop (1987)	Contingent valuation
Maximum range: $126	Wolf, endangered species in Sweden Boman and Bosdedt (1995)	Contingent valuation
Multiple species		
Minimum range: $18	Preservation of threatened and endangered species populations in the US, Hageman (1985)	Contingent valuation
Maximum range: $194	Preservation of 300 endangered species in Sweden, Johnansson (1989)	Contingent valuation
Habitat: Terrestrial (non-use)		
Minimum range: $27	Protection of the Nadgee Nature Reserve, Australia, Bennett (1984)	Contingent valuation
Maximum range: $101	Desert Protection in California, US Richer (1995)	Contingent valuation
Habitat: Coastal (non-use)		
Minimum range: $10	Protection of New Jersey beaches, US Silberman *et al.* (1992)	Contingent valuation
Maximum range: $51	Protection of a wilderness coastal area, Portugal Nunes (2000b)	Contingent valuation
Habitat: Wetland (non-use)		
Minimum range: $8	Protection of the Norfolk Broads, UK Batemann *et al.* (1992)	Contingent valuation
Maximum range: $96	Enhancing wetland habitat in California, US Hoehn and Loomis (1993)	Contingent valuation
Habitat: Ecosystem space (recreation)		
Minimum range: $23/trip	Forest recreation activities in Flanders, Belgium Moons (1999)	Travel cost
Maximum range: $23 million/year	Tourism in Ecuador WTO (1997)	Tourism revenue
Ecosystems functions		
Minimum range: $1.2 million	Life-support value of a wetland ecosystem in the a Swedish island, Baltic Sea, Turner *et al.* (1995)	Production function
Maximum range: $4.4 billion	Water ecosystem benefits in ten regions in US Ribaudo (1989)	Averting behaviour

studies that focus on ecosystems functions and the value assessment of life-support, soil and wind erosion, or water quality benefits. Some of these studies are listed in Table 6.

During the 1980s many contingent valuation studies dealt with the measurement of the non-use benefits derived from the conservation of terrestrial parks and nature reserves - see Bennett (1984) and Richer (1995). The valuation applications continued through the 1990s –

see Silberman *et al.* (1992), Batemann *et al.* (1992), and Hoehn and Loomis (1993) – but now also tackling the valuation of non-use benefits of coastal and wetland habitats. Silberman *et al.* estimated the existence value for users and nonusers of New Jersey beaches. The results show that the mean WTP for a user is about $15.1, while the mean WTP for a nonuser is about $9.26. Batemann *et al.* (1992) undertook a contingent valuation to assess the monetary value of conserving the Norfolk Broads, a wetland site in the UK gathering three National Nature Reserves. A mail survey across Britain showed that non-visitors respondents were willing to pay, on average, 4 pounds (circa $8). More recently, Nunes (2000b) used for the first time the CV method in Portugal to assess the national WTP for the protection of a coastal natural area. The mean WTP results ranged from $40 to $51.

In the recreation domain, the World Tourism Organisation (WTO 1997) estimated that Ecuador earned $255 million from eco-tourism in 1995. A major sum accrued to a single park, the Galapagos Islands. Studies of less popular areas indicate lower values. The recreational value of a Regional Forest Park in Belgium was estimated to be around $23 per trip (Moons 1999). Norton and Southey (1995) calculated the economic value of biodiversity protection in Kenya by assessing the associated opportunity costs in terms of forgone agricultural production, which is estimated to be $203 million. This valued can be compared with $42 million in net financial revenues from wildlife tourism and forestry. More recently, Chase *et al.* (1998) studied the eco-tourism demand in Costa Rica. The value estimates result from the survey of foreign visitors to three national parks: Volcan Irazu, Volcan Poas, and Manuel Antonio. The highest WTP registered was about $25 for the Manuel Antonio national park.

When it comes to the monetary valuation of ecosystem functions, CV may not be the first method choice. This is because ecosystem life support is not an issue that the general population is familiar with. In addition, the complexity of the relationships involved makes an accurate and comprehensive survey description harder. Researchers frequently end up with the use of valuation methods such averting behaviour, production function, or hedonic pricing. In 1991 Andreasson-Gren (1991) estimated the costs of nitrogen abatement via wetlands restoration with the costs for using conventional technologies. The estimated nitrogen purification capacity of wetlands was based on the results of a Swedish island in the Baltic Sea, Gotland. According to the study results, the total value of a marginal increase in nitrogen abatement on Gotland was about SEK 968 per kilogram. Turner *et al*. (1995) addressed the valuation of a wetland ecosystem in the Swedish island in the Baltic Sea exploring the use of production function. Their value estimations confirmed that a considerable amount of industrial energy would be necessary to substitute for the loss of wetland life-support functions – see results in Table 6. Ribaudo (1989) is responsible for one of the most comprehensive study valuing water ecosystems. The author valued the economic benefits from the reduction in the discharge of pollutants in waterway systems for nine impact categories: recreational fishing, navigation, water storage, irrigation ditches, water treatment, industrial water use, steam cooling, and flooding. Benefits were defined in terms of changes in defensive expenditures, changes in production costs, or changes in consumer surplus, depending on the damage category and the availability of data. The total water quality benefits were estimated to be $4.4 billion.

4. Integrated ecological-economic modelling and valuation of biodiversity

4.1 Economics – Ecology interface

The analysis and the modelling of biodiversity are rooted in both natural and social science domains and thus imply the study of human economic activities, its relationships with respect to biodiversity, the structure and functions of ecosystems. The combination or integration of the two approaches will imply a somewhat qualitative, formal, sequentially integrated framework. Interdisciplinary work involves economists or ecologists transferring elements or even theories and models from one discipline to another and transforming them for their specific purpose (Perrings *et al.* 1995). The underlying objective of this approach is the development of a common way of thinking about modelling and valuation of biodiversity. For instance, if economic and ecological models fit in a (general) systems frame, then they may be blended in a single model structure, where compartments or modules may represent the original models, and certain outputs of one module serve as input for another. Nevertheless, it is often not easy to link models directly. Alternatively, if both the economic and ecological systems are represented in the form of programming or optimisation models then several options are available: look for a new, aggregate objective; adopt a multi-objective or conflict analysis framework; or, when possible, derive multiple sets of optimality conditions and solve these simultaneously. Finally, when economic and ecological systems are represented by different model types, it is harder to suggest how they can be linked to one another. Where economic models have an optimisation format and ecosystem models have a descriptive format, then a direct technical integration seems feasible, otherwise heuristic approaches are needed. This may require operations such as reduction, simplifying or summarising. For example, one can come up with a simple dynamic model summarising and simplifying some of the causal relationships of the spatial hydrological model and the statistical vegetation model, and linking the outcomes to a simplified economic interaction and values model.

Before discussing specific methods and models it seems useful to say a few words about pros and cons of integration frameworks and respective conceptual perspectives. These are discussed in the following section.

4.2 Frameworks and theories underlying integrated modelling

A general method to develop integrated models is a systems approach (also ‘systems dynamics’). This covers a wide range of model types: linear versus nonlinear, continuous versus discrete, deterministic versus stochastic, and optimising versus descriptive. Such system approaches allow dealing with concepts like ecological dynamic processes, feedback mechanisms, and controlling strategies (see Bennett and Chorley 1978; Costanza *et al.* 1993). One can integrate two subsystems, or have a hierarchy or nesting of systems. The fixed elements in the system can be either considered as black boxes or described as empirical or logical processes themselves. The systems approach is suitable to integrate existing models, and to incorporate temporal as well as spatial processes. Costanza *et al.* distinguish economic, ecological and integrated approaches on basis of the criteria: (1) generality, characterised by simple theoretical or conceptual models that aggregate, caricature and exaggerate; (2) precision, characterised by statistical, short-term, partial, static or linear models with one element examined in much detail; and (3) realism, characterised by causal, non-linear, dynamic-evolutionary, and complex models. These three criteria are usually conflicting, so

that a trade-off between them is inevitable. A very general and almost non-theoretical ('no assumptions') framework is the Driver-Pressure-State-Impact-Response (DPSIR) framework, a variation on the framework proposed for environmental data classification by Turner *et al.* (1999) and Rotmans and de Vries (1997) for integrated analysis and modelling. The components have the following interpretation:

- 'Driver' = economic and social activities and processes;
- 'Pressure' = pressures on the human (health) and environmental system (resources and ecosystems);
- 'State' = the physical, chemical and biological changes in the biosphere, human population, resources and artefacts (buildings, infrastructure, machines);
- 'Impact' = the social, economic and ecological impacts of natural or human-induced changes in the biosphere;
- 'Responses' = human interventions on the level of drivers (prevention, changing behavior), pressures (mitigation), states (relocation) or impacts (restoration, health care).

According to Rotmans and de Vries (1997) integration can include various types. Vertical integration means that the causal chain in the PSIR or DPSIR framework is completely described in a model ("close the PSIR loop"; p. 25). Horizontal integration (of subsystems) in this context is defined as the coupling of various global biogeochemical cycles and earth system compartments (atmosphere, terrestrial biosphere, hydrosphere, lithosphere and cryosphere). An alternative and relevant distinction is between analytical and heuristic integration. Analytical integration means combining all aspects studied in a single model (and therefore model type). Heuristic integration can proceed by using the output of one model as input to another, and vice versa, as well as extending this by an (finite) iterative interaction. In this case different model types can be combined, such as optimisation models and descriptive models. If one desires to attain a great deal of analytical power, the analytical integration seems attractive, whereas striving for realism would imply the use of a heuristically linked set of models from different disciplines.

Formal modelling and evaluation in integrated economic-environmental studies has both advantages and disadvantages. Three main advantages are: (1) handling data, information, theories, and empirical findings from various contributing disciplines in a systematic and consistent way; (2) being explicit about assumptions, theories and facts; and (3) addressing complex phenomena, interactions, feedback, laborious calculations and temporally, spatially and sectorally detailed and disaggregate processes. Most disadvantages of integrated modelling apply to non-model based integrated research as well.[16] They include: unclear synergy of approximations and uncertainties; a rough application of monodisciplinary theories and empirical insights; a simplification of complex phenomena (e.g., by treating them as a black box); a misinterpretation and arbitrary choice of disciplinary perspectives by the model, and a lack of systematic or complete linking of subsystems or submodels. Complex or high-dimensional models have the extra disadvantage of being difficult to calibrate and validate, and of lacking transparency.

[16] An argument against non-formal approaches to integrated research is that these fail to provide for a systematic and consistent linking of data, theories and empirical insights from various disciplines. Instead, these approaches tend to result in a battle of perspectives based on distinct and usually implicit premises and information bases. Models force researchers at least to be explicit about the latter two inputs to integrated research.

The main disadvantage of models perhaps is that they are trusted too much, that they run the risk of being interpreted as objective representations of reality, and then are taken too seriously by especially laypersons and policy makers. On the other hand, policy makers often indicate their doubts about formal modelling. Shackley (1997) states that numerical models have, despite their long tradition of development and widespread use, not achieved the epistemological status that the controlled laboratory experiment has in natural sciences (and more recently in social sciences and environmental economics in particular; see Shogren 1999). This relates to the fact that modelling results never 'prove' anything, since they do not generate real or physical processes. The best way to view theoretical and especially empirical models is to consider them as tools for hypothetical experiments with complex systems, which serve as analogies or pictures of real-world systems that do not allow – technically, morally or politically – for experimentation. In other words, complex model systems, notably integrated economic-ecological models, are heuristic devices for learning about the real-world system, rather than for predicting its real course. In addition, integrated modelling is also restricted by the model type.

4.3 Integrated assessment

Full or total economic-environmental integration means a combination, leading to the complex linking of various drivers, pressures, states, impacts and responses, thus allowing for various synergies and feedback. The literature shows various examples of such integrated economic-environmental frameworks. Surveys are offered by Barbier (1990), van den Bergh and Nijkamp (1991), van den Bergh (1996), Costanza *et al.* (1997), Ayres *et al.* (1999) and Turner *et al.* (1999).

Many integrated models defined at the level of ecosystems are based on the standard systems-ecological approach (Patten 1971; Jørgenson 1992). They include ecosystem modules that describe the effects of environmental pollution, resource use and other types of disturbance. A main problem is modelling the effects of multiple stress factors, as the empirical basis for this is often lacking. Various integrated models have been developed for terrestrial and aquatic systems. Surveys are presented in Braat and van Lierop (1987), van den Bergh (1996) and Costanza *et al.* (1997). Some studies have paid much attention to spatial aspects, focusing on spatial disaggregation into zones (for instance, Giaoutzi and Nijkamp, 1993; van den Bergh and Nijkamp 1994) or land-use planning in interaction with landscape ecology (see Bockstael *et al.* 1995). Formal theoretical approaches in ecology that provide a basis for these approaches have been described by Watt (1968), Maynard-Smith (1974), Roughgarden *et al.* (1989) and Jørgenson (1992). Perrings and Walker (1997) consider resilience in a simple integrated model of fire occurrences in semi-arid rangelands such as found in Australia. The model describes the interaction between extreme events (fire, flood, and draughts), grazing pressure, and multiple locally-stable states. Carpenter *et al.* (1999) develop and explore water and land use options in an integrated model of a prototypical region with a lake that is being polluted. This model combines bounded rationally behaviour, supposedly in accordance with the reality in regional resource and environmental management, and a nonlinear ecosystem module describing processes occurring at different speeds. The model generates multiple locally stable states as well as "flipping" behaviour (see also Janssen *et al.* 1999). Swallow (1994) integrates theoretical models of renewable and non-renewable resources to address multiple use and tradeoffs in wetland systems.

A special category of integrated modelling is sometimes referred to as the biophysical or energy approach. This aims to integrate economic and environmental, ecological processes in energy-physical dimensions, based on the notion that any system is constrained by energy availability (Odum, 1983). These models include energy and mass balances. A central concept in this approach is 'embodied energy', defined as the direct and indirect energy required to produce organised material structures. Applications of these energy-inspired models cover ecosystems, economic systems, and environment-economy models (Odum, 1987). An extended application to a regional system is Jansson and Zuchetto (1978) (see also Zuchetto and Jansson 1985).Biophysical assessment models, and other integrated empirical applications, are reviewed and discussed in the following section.

4.4 Empirical applications

Striving for empirically sound models often implies modest approaches to improve precision, which usually goes at the cost of model use in a wider context or wider range of parameter values. The development of integrated models, by joint efforts of economists and ecologists, is based on bringing together elements, theories or models from each discipline and transforming these for the purpose of integration. This may require steps like reduction, simplifying or summarising. The results may not always be greeted with enthusiasm within the disciplines, especially when they neglect certain nuances or different viewpoints. For example, the recent focus on integrated assessment of the enhanced greenhouse effect (potential climate change) can be regarded as the new wave in 'World models', where (again) economists and others have tended to rely on different model approaches (Bruce *et al.* 1995). The integrated climate assessment models provide AN integration of natural sciences (physics, chemistry, biology, earth sciences) and social sciences (economics, sociology, political science), and have so far given rise to a continuation of the trend in World models towards increasing detail and disagregation. These climate assessment models have a multilayered conceptual structure that distinguishes physical and environmental effects of human activities from adjustments to climate change by humans (individuals, firms, organizations) and policy responses (mitigation, aimed at the causes) at various spatial levels (Parry and Carter 1998) – see Table 7. As we can see, integrated models can have different formats. One important distinction is between policy optimization and evaluation (usually numerical simulation) models. One of the first and famous integrated assessment models used in policy making is the model RAINS (Alcamo *et al.* 1990). This includes an optimisation algorithm for calculating cost-effective acidification strategies in Europe, aimed at realising deposition targets throughout Europe, and taking account of sensitive natural areas (forests and lakes). This model is a rare case of direct science-policy influence, as it was used in the negotiations on transboundary air pollution in Europe. Castells (1999) offers an informative analysis of the institutional and evolutionary dimensions of the interaction between scientists, research institutions and negotiations on international environmental agreements, with special attention given to the RAINS model and acid rain context in Europe.

In the area of integrated assessment models for CO_2 emissions (climate) strategies, one can find both economic optimization (Nordhaus 1994) and detailed descriptive model systems like IMAGE and TARGETS (Alcamo 1994, and Rotmans and de Vries 1997). DICE by Nordhaus (1991) is the first example of a policy optimization model for climate change. The model essentially combines economic growth theory with a simplified climate change model. Tol (1998) provides a short account of the evolution of the economic optimization approach

Table 7: Characterising integrated models

Model criterion	Range of choice	Examples of distinct approaches
Analytical integration	Optimisation (benevolent decision maker); Equilibrium (partial or general); Game-theoretical; Dynamic-mechanistic; Adaptive (multi-agent & dynamic); Evolutionary (irreversible, bounded rationality)	Many theoretical models: growth theory, renewable resource economics (fisheries, forestry, water quality/quantity), systems models (Limits-to-growth, Meadows), cost-effectiveness models (RAINS), welfare optimization (DICE)
Heuristic integration	Satellite principle; Multilayer Subsystems; Sequential; Parallel consistent scenarios; Aggregation of Indicators; Evaluation	regional environmental quality models (Resources for the Future), World models (Club of Rome), integrated assessment model, the present study
Spatial coverage	World; National; Regional; Urban, Local, Ecosystem	Ecosystem modelling, macroeconomic modelling, regional modelling, urban modelling, world models
Spatial disaggregation	Single region; Multi-region; Spatial grid (GIS)	Integrated assessment models (Climate change), land use models
Aggregation level	Micro (indiduals, households); Macro (national economy, main sectors, global); Sectoral; Interest groups; Homogeneous land plots; Spatial grids; Temporal (days, seasons, years)	Computable general equilibrium models, macroeconomic models (Keynesian), multisector models, land use models, landscape models

Source: van den Bergh (2000)

to climate change research. He emphasizes the attention for analysis of uncertainty and learning from a cost-effectiveness perspective, which has given rise to various model formulations and analyses. More recently, Janssen (1998) and van Ierland (1999) present informative surveys and categorizations of macroeconomic-cum-environment and macro-level integrated models, including the climate oriented integrated assessment models. Van Ierland devotes special attention to the various "regionalized world" models (with acronyms like RICE, CETA, MERGE, DIALOGUE, FUND). Van den Bergh and Hofkes (1998) have collected distinct approaches to integrated models with an economic emphasis focusing on sustainable development questions, in theory and practice as well as at global and regional levels.

4.5 Integrating modelling and monetary valuation

Progress on improving methods for providing such economic information (particularly predictive information) will require a strong and dynamic interdisciplinary dialogue. At this level integrating modelling and monetary valuation can present important advantages for policy guidance, presenting important interactions. First, values estimated in a valuation study can be used as a parameter value in a model study. Benefits or value transfer (e.g. meta-analysis exercise) can be used to translate value estimates to other contexts, conditions, locations or temporal settings that do not allow for direct valuation in 'primary studies' (due to technical or financial constraints). Second, models can be used to generate values under

particular scenarios. In particular, dynamic models can be used to generate a flow of benefits over time and to compute the net present value, which can serve as a value relating to a particular scenario of ecosystem change or management.

Third, models can be used to generate detailed scenarios that enter valuation experiments. An input scenario can describe a general environmental change, regional development or ecosystem management. This can be fed into a model calculation, which in turn provides an output scenario with more detailed spatial or temporal information. The latter can then serve, for example, as a hypothetical scenario for valuation, which is presented to respondents in a certain format (graphs, tables, story, diagrams, pictures) so as to inform them about potential consequences of the general policy or exogenous change. Computer software can be used in such a process. Finally, the output of model and valuation studies can be compared. For instance, when studying a scenario for wetland transformation one can model the consequences in multiple dimensions (physical, ecological and costs/benefits), and aggregate these via a multi-criteria evaluation procedure, with weights being set by a decision-maker or a representative panel of stakeholders. Alternatively, one can ask respondents to provide value estimates, such as a willingness to pay for not experiencing the change. If such information is available for multiple management scenarios, then rankings based on either approach can be compared.

5. Conclusions and recommendations

How can we use the ideas presented in Sections 2, 3 and 4 to formulate an integrated, effective framework addressing to the economic valuation of biodiversity? And what do we learn from the empirical valuation studies focused on biodiversity? A crucial, and initial step, is to carefully describe the object of analysis and valuation. Therefore, researchers face two important decisions when valuing biodiversity: (1) at which level should biodiversity be examined; and, (2) what biodiversity value types should be measured?

Assessment of the physical interactions

Clearly, it is necessary to attain information about the nature, type, and persistence of stress or shocks experienced by ecosystems, their functions and stability, and their impacts on human welfare (loss). A comprehensive assessment of ecosystem biodiversity characteristics, structure and functioning requires the analyst to undertake various important actions. First, the causes and consequences of biodiversity degradation/loss should be determined in order to improve understanding of socio-economic impacts on biodiversity processes and attributes. Second, the sustainability of biodiversity uses and negative impacts on the biodiversity caused by off-site human activities should be assessed. the range and degree of biodiversity functioning should be assessed, especially in terms of ecosystem-functional relationships. Third, the range and degree of biodiversity functioning should be assessed, especially in terms of ecosystem-functional relationships. Finally, alternative biodiversity conversion, development and conservation management strategies should be assessed and spatial and temporal systems analysis (via a range of methods and techniques) of alternative biodiversity change scenarios should be carried out.

Ecological evaluation

The physical assessment of the functions performed by biodiversity is an essential prerequisite to any ecological evaluation. However, simply identifying these functions is

insufficient, if we want to present resource managers and policy makers with the relevant policy response options. It is necessary to develop criteria for the expression of the functions in a form that allows evaluation. Finally, values are attached to particular levels of criteria. The provision of such information is essential if an efficient level of biodiversity resource conservation is to be determined.

Ecologists contribute to the identify the range of policy decisions with respect to the biodiversity management/conservation strategies, by exploring the use of ecological valuation methods such as red data species lists and biological value indexes. More recently, computer models have been often used by ecologists to aid conservation evaluation decisions, namely in the domain of species population. Several general models have been developed for this task (e.g., VORTEX and METAPOP) and applied to calculate minimum dynamic areas, i.e. the geographic area of suitable habitat required to support the minimum viable population (e.g., the giant panda). Finally, computer models have also been used for habitat evaluation with the objective of simulating the impact of different development scenarios and predicting the respective ecological conservation values. This approach to ecological evaluation has attractive features. First, it permits the introduction of evaluation criteria within a transparent and repeatable framework. Second, it allows a direct comparison of management, or conservation, strategies. Finally, it permits an evaluation even when some of the data concerning the functions performed by biodiversity is are missing, which is quite common in practice.

Economic valuation

The concept of economic value has its foundations in welfare economics. It focuses on the welfare of humans, and in turn depends on the theory of consumer behaviour. Therefore, valuation in an economic sense is always the result of an interaction between the subject and an object. Moreover, economists do not pursue total value assessment of an environmental system, but rather system changes. This means that the terms ‘economic value’ and ‘welfare change’ can be used interchangeably. The goal is then to assess the human welfare significance of biodiversity changes, through determination of the changes in the provision of biodiversity related goods and services and consequent impacts on the well-being of humans who derive use or non-use benefits from such a provision. Biodiversity has been, however, frequently tackled in the empirical valuation studies with the measurement of the economic value of a particular set of biological resources, such as changes of the range of species on human recreation activities. Though, biodiversity has indirect links with various ecosystem goods and services. One reason why this monetary value component is not frequently assessed relates to the complexity involved in the ecosystems and thus their functions and processes can not be easily characterised. Therefore, one needs to have these two concepts clear and do not confuse the biodiversity economic valuation exercise with the economic value (of a specific) biological resource.

Different instruments are available to assess the economic value of biodiversity. Survey valuation studies have often been used because the use of revealed preference methods will leave out important biodiversity value types, notably non-use and quasi-option values. This can lead to a significant value measurement bias, especially if species conservation decisions are characterised by a high irreversibility. Alternatively, researchers can combine valuation techniques. Special attention, however, should then be given to value aggregation across the resulting values so as to avoid double counting.

Integrating modelling

The analyses and valuation of biodiversity are rooted in both natural and social science domains and thus imply the study of human economic activities, its relationships with respect to biodiversity, and the structure and functions of ecosystems. In practice, there is a clear need to obtain information about the cause, type, and persistence of stress on biodiversity and the estimation of the respective impacts on human welfare. The combination or integration of the ecological and ecological to assess and value biodiversity leads to an integrated framework. The underlying objective of this approach is the development of a common way of thinking about modelling and valuation of biodiversity. Interdisciplinary work is therefore required, involving economists or ecologists transferring elements or even theories and models from one discipline to another and transforming them for their specific purpose. This may require analytical operations such as reduction, simplifying or summarising. For instance, one can come up with a simple dynamic model by simplifying some of the causal relationships of the spatial hydrological model and the statistical vegetation model, and linking the outcomes to a simplified economic interaction and values model.

Last but not least, economists need to be aware of the limitations of any proposed valuation exercise. Despite all research efforts in an integrated, multidisciplinary analysis, modelling and valuation, it should be recognised that not all biodiversity value types can be made explicit nor that all explicit biodiversity value types can be measured in monetary terms. A valuation study is a partial and characterised by strict temporal and spatial boundary. As Gowdy (1997) has recently said "… although values of environmental services may be used to justify biodiversity protection measures, it must be stressed that value constitutes a small portion of the total biodiversity value…". In other words, monetary estimates of biodiversity should at best be interpreted as conservative estimations and thus regarded in terms of lower bounds. Clearly, more valuation studies need to be pursued to improve the information available to decision makers about the nature, magnitude and value on the impacts of human activities on the overall level of biodiversity.

Annex

The Criteria for Critically Endangered, Endangered and Vulnerable. Prepared by the IUCN Species Survival Commission As approved by the 40th meeting of the IUCN Council, Gland, Switzerland 30 November 1994.

Critically Endangered (Endangered/ Vulnerable)

A) Population reduction in the form of either of the following:

1. An observed, estimated, inferred or suspected reduction of at least 80% (50%/20%) over the last 10 years or three generations, whichever is the longer, based on (and specifying) any of the following:
    (a) direct observation
    (b) an index of abundance appropriate for the taxon
    (c) a decline in area of occupancy, extent of occurrence and/or quality of habitat
    (d) actual or potential levels of exploitation
    (e) the effects of introduced taxa, hybridisation, pathogens, pollutants, or parasites.
2. A reduction of at least 80%, (50%/20%) projected or suspected to be met within the next ten years or three generations, whichever is the longer, based on (and specifying) any of (b), (c), (d) or (e) above.

B) Extent of occurrence estimated to be less than 100 km2 (5,000 km2/ 20,000 km2) or area of occupancy estimated to be less than 10 km2, and estimates indicating any two of the following:

1. Severely fragmented or known to exist at only a single location.
2. Continuing decline, observed, inferred or projected, in any of the following:
    (a) extent of occurrence
    (b) area of occupancy
    (c) area, extent and/or quality of habitat
    (d) number of locations or subpopulations
    (e) number of mature individuals.
3. Extreme fluctuations in any of the following:
    (a) extent of occurrence
    (b) area of occupancy
    (c) number of locations or subpopulations
    (d) number of mature individuals.

C. Population estimated to number less than 250 (2,500/10,000) mature individuals and either:

1. An estimated continuing decline of at least 25% (20%/10%) within 3 years or one generation (5 years or two generations/10 years or three generations), whichever is longer, or
2. A continuing decline, observed, projected, or inferred, in numbers of mature individuals and population structure in the form of either:
    (a) severely fragmented (i.e. no subpopulation estimated to contain more than 50 mature individuals)
    (b) all individuals are in a single subpopulation.

D. Population estimated to number less than 50 (250/1,000) mature individuals.

E. Quantitative analysis showing the probability of extinction in the wild is at least 50% (20%/10%) within 10 years or 3 generations (20 years or 5 generations/100 years), whichever is the longer.

References

Akcakaya, H.R. (1994) "RAMAS/METAPOP: viability analysis for stage-structured metapopulations, *Applied Biomathematics*, Setauket, New York, US.

Alcamo, J., R. Shaw and L. Hordijk (eds.), 1990. *The Rains Model of Acidification: Science and Strategies in Europe*. Kluwer Academic Publishers, Dordrecht.

Alcamo, J. (ed.), (1994) *Image 2.0: Integrated Modelling of Global Climate Change*. Kluwer Academic Publishers, Dordrecht.

Andreasson-Gren, I.M. (1991) "Costs for Nitrogen Source Reduction in a Eutrophicated Bay in Sweden" in Folke C. and T. Kraberger (eds), *Linking the Natural Environment and the Economy: Essays from the Eco-group*, Kluwer Academic Publishers.

Antonovic (1990) "Genetically Based Measures of Uniqueness", in Orians, G.H., Brown, G.M., Kunin, W.E. and J.E. Swierzinbinski (eds) *The Preservation and Valuation of Genetic Resources*,University of Washington Press, pp. 94-118.

Ayres, R.U., K.J. Button and P. Nijkamp (1999) *Global Aspects of the Environment*, Vol. 1 and 2. Edward Elgar, Cheltenham, UK.

Barbier, E.B. (1990) "Alternative approaches to economic-environmental interactions", *Ecological Economics*, 2, pp. 7-26.

Barbier, E.B. (1994) "Valuing Environmental Functions: Tropical Wetlands", *Land Economics*, 70(2), pp. 155-173

Batemann, I.J., Willis, K.G., Garrod, G.D., Doktor, P., Langford I. and R.K. Turner (1992) *Recreation and Environmental Preservation Value of the Norfolk Broads: A Contingent Valuation Study*, Report to the National Rivers Authority, Environmental Appraisal Group, University of East Anglia, valuation results cited in Turner et al. (1995).

Becker, G.S. (1993) "Nobel Lecture: the Economic Way of Looking at Behaviour", *Journal of Political Economy*, 101(3), pp. 385-409.

Bennett, J. W. (1984) "Using Direct Questioning to Value the Existence Benefits of Preserved Natural Areas", *Australian Journal of Agricultural Economics*, 28, pp. 136-152.

Bennett, R.J. and R.J. Chorley, 1978. *Environmental Systems*. Methuen & Co., London.

Bertollo, P. (1998) "Assessing Ecosystem Health in Governed Landscapes: a Framework for Developing Core *Ecosystem Health*, 4(1), pp. 33-51.

Bockstael, N. E., McConnell, K. E., and I. Strand (1991) "Recreation", in Braden, J. B. and C. D. Kolstad (eds) *"Measuring the Demand for Environment Quality"*, Chapter VIII, Elsevier Science Publishers, North-Holland.

Bockstael, N., R. Costanza, I. Strand, W. Boynton, K. Bell, and L. Waigner (1995) "Ecological Economic modeling and valuation of ecosystems", *Ecological Economics*, 14, pp. 143-159.

Boman, M. and G. Bostedt (1995) "Valuing the Wolf in Sweden", *Report No. 110, Swedish University of Agricultural Sciences*, Department of Forest Economics, Umeä.

Boon, P.J., Holmes, N.T, Maitland, P.S., Rowell, T.A. and J. Davies (1997) "A system for evaluating reivers conservation (SERCON): development, structure and function", in P.J. Boon and D.L. Howell (eds) *Freshwater Quality: Defining the Indefinable*, HMSO, Edinburgh, UK.

Boyle, K.J. and R.C. Bishop (1987) "Valuing wildlife in benefit-cost analysis: a case study involving *Water Resources Research*, 23, pp. 943-950.

Braden, J. B., Kolstad, C.D. and D. Miltz (eds) (1991) "Measuring the Demand for Environment Quality", Elsevier Science Publishers, North-Holland.

Braat, L.C. and W.F.J. van Lierop (eds.) (1987) *Economic-Ecological Modelling*. North- Holland.

Brink, B.J and S.H. Hosper (1989) "Naar Toetsbare Ecoloische Doelstllingen voor het Waterbeheer: de AMOEBE benadering", *H_2O*, 22(20), pp. 607-617.

Bruce, J., *et al.* (1995) *Climate Change 1995: Economic and Social Dimensions of Climate Change*, Contribution of Working Group III, Cambridge University Press.

Buys, S. (1995) "Urbanisation Alternatives in the Fourth Report on Physical Planning", *Netherlands Journal of Housing and the Built Environment"*, 7(2), pp. 135-154.

Carpenter, S., W. Brock, and P. Hanson (1999) "Ecological and social dynamics in simple models of ecosystem *Conservation Ecology* 3(2): 4-12.

Carson, R. T., R. C. Mitchell, W. M. Hanemann, R. J. Kopp, S. Presser and P. A. Ruud (1992) "A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill", *Report prepared for the Attorney General of the State of Alaska*, Washington.

Castells Cabré, N. (1999) *International Environmental Agreements: Institutional Innovation in European Transboundary Air Pollution Policies*. Ph.D. dissertation, Free University, Amsterdam.

Chase, L.C., Lee, D.R., Schulze, W.D. and D.J. Anderson (1998) "Ecotourism Demand and Differential Pricing of National Park Access in Costa Rica", *Land Economics*, 74(4), pp. 466-482.

Cody, M.L. (1986) "Diversity , Rarity and Conservation in Mediterranean-Climate Regions", in M.E. Soule (ed) *Conservation Biology: the Science of Scarcity and Diversity*, pp. 123-152, Sinauer Associates, Sunderland, Mass.

Cody, M.L. (1993) "Bird Diversity Components within and between Habitats in Australia, in R.E. Ricklefs and D. Schluter (eds), *Species Diversity in Ecological Communities: Historical and Geographic Perspectives*, pp. 147-158, University of Chicago Press, Chicago, US.

Costanza, R. (1992) "Toward an Operational Definition of Ecosystem Health", in Costanza, R., Norton, B. and B. Haskell (eds), *Ecosystem Health: New Goals for Environmental Management*, Island Press, Washington, US.

Costanza, R., Norton, B. and B. Haskell (eds) (1992) "Ecosystem Health: New Goals for Environmental Management", Island Press, Washington, US.

Costanza, R., L. Wainger, C. Folke and K.-G. Mäler (1993) "Modeling complex ecological economic systems", *BioScience*, Vol. 43: 545-555.

Costanza, R., Kemp, M. and W. Boyton (1995) "Scale and Biodiversity in Coastal and Estuarine Ecosystems" in Perrings, C., C.S. Holling, K-G. Maler, B-O. Jansson and C. Folke (eds) *Biodiversity Loss: Economic and Ecological Issues*, Cambridge University Press.

Costanza, R., C. Perrings, and C.J. Cleveland (eds.) (1997) *The Development of Ecological Economics*, Edward Elgar, Cheltenham, UK.

Cropper, M. L. and A. M. Freeman III (1991) "Environmental Health Effects", in Braden, J. B. and C. D. Kolstad (eds) "*Measuring the Demand for Environment Quality*", Chapter III, Elsevier Science Publishers, North-Holland.

Cuperus, R. and K.J. Canters (1995) "Op naar natuurcompositie voor autosnelwegen", Rijksuniversiteir Leiden, Leiden.

de Groot, R.S. (1994) "Environmental functions and the economic value of natural ecosystems", in *Investing in Natural Capital: the Ecological Economics Approach to Sustainability*, A. Jansson *et al.* (eds), Island Press.

Ehrenfeld, D. (1988) "Why put a value in biodiversity" in *Biodiversity*, E.O. Wilson (ed), National Academy Press.

Fitter, R., and M. Fitter (ed.) (1987) *The Road to Extinction*, Gland, Switzerland: IUCN.

Folke C. and T. Kraberger (eds) (1991) *Linking the Natural Environment and the Economy: Essays from the Eco-group*, Kluwer Academic Publishers.

Folke, C., C.S. Holling and C. Perrings (1996) "Biological diversity, ecosystems and the human scale", *Ecological Applications*, 6, pp. 1018-1024.

Freeman, A. M. III (1979) "Approaches to Measuring Public Goods Demands", *American Journal of Agricultural Economics*, 61, pp. 915-920.

Franklin, J.F. (1993) "Preserving Biodiversity: species, ecosystems or landscapes", *Ecological Applications*, 3, pp. 202-205.

Giaoutzi, M. and P. Nijkamp (1993) *Decisions Support Models for Regional Sustainable Development.* Avebury, Aldershot.

Gowdy, J.M. (1997) "The Value of Biodiversity: Markets, Society and Ecosystems", *Land Econmics*, 73(1), pp. 25-41.

Gren, I-M., Folke, C., Turner, R.K. and Bateman, I. (1994) Primary and Secondary Values of Wetland Ecosystems. *Environmental and Resource Economics*. Vol.4: pp. 55-74.

Hageman, R. (1985) "Valuing Marine Mammal Populations: Benefits Valuations in Multi-species Ecosystems", in *Administrative Report*, Southwest Fisheries, values cited in Eagle J. and D. Betters (1998) "The Endangered Species Act and Economic Values: a Comparison of Fines and Contingent Valuation Studies", *Ecological Economics*, 26, pp. 165-171.

Hanley, N., C. L. Spash and L. Walter (1995) "Problems in Valuing the Benefits of Biodiversity Protection", *Environment and Resources Economics*, 5(3), pp. 249-272.

Hausman, J. A. (ed.) (1993), "Contingent valuation: A critical assessment", Contributions to Economic Analysis, North-Holland, New York, US.

Hoehn, J.P. and J.B. Loomis (1993) "Substitution Effects in the Valuation of Multiple Environmental *Journal of Environmental Economics and Management*, 25(1), pp. 56-75.

Holling, C.S. (1987) Simplifying the complex, the paradigms of ecological function and structure", *European Journal of Operational Research*, 30, pp. 139-146.

Holling, C.S. (1992) "Cross-scale morphology geometry and dynamics of the ecosystems", *Ecological Monographs*, 62, pp. 447-502.

Howels, O. and Edward-Jones (1997) "A feasibility study of reintroducing wild boar to Scotland: would existing *Biological Conservation*, 81, pp. 77-89.

Huston, M.A. (1994) *Biological Diversity – the coexistence of species on changing landscapes,* Cambridge University Press, Cambridge, UK.

Huszar, P.C. (1989) "Economics of Reducing Off-Site Costs of Wind Erosion", *Land Economics*, 65(4), pp. 333-340.

ICN-Instituto de Conservação da Natureza (1993) "Livro vermelho dos vertebrados em Portugal", Secretatia de Estado do Ambiente e Consumidor, ICN, Lisboa, Portugal.

Ierland, E.C. (1999) Environment in macroeconomic modelling. In: J.C.J.M. van den Bergh (ed.). *Handbook of Environmental and Resource Economics*. Edward Elgar, Cheltenham.

IUCN (1991) "United Nation Environmental Program and World Wide Fund for Nature, Caring for the Earth", International Union of Nature Conservation, Gland.

IUCN (1993) *Draft IUCN Red List Categories,* Gland, Switzerland: IUCN.

IUCN (2000) *IUCN Internet Page,* http://www.iucn.org/themes/ssc/redlists/categor.htm

Janssen, M.A., B.H. Walker, J. Langridge and N. Abel (1999) "An integrated model of rangelands as a complex adaptive system", *Working paper,* CSIRO Wildlifle and Ecology, Canberra, Australia.

Janssen, M.A. (1998). *Modelling Global Change: The Art of Integrated Assessment*. Modelling. Edward Elgar Publishers, Cheltenham.

Jansson, AM. and J. Zucchetto (1978) *Energy, Economic and Ecological Relationships for Gotland, Sweden: A Regional Systems Study*. Ecological Bulletins 28. Swedish Natural Science Research Council, Stockholm.

Johnansson, P-O. (1989) "Valuing Public Goods in a Risky World: an Experiment", in Folmer, H. and E. Ierland (eds) *Evaluation Methods and Policy Making in Environmental Economics*, Amsterdam, North Holland.

Jørgenson, S.E. (1992) *Integration of Ecosystem Theories: A Pattern.* Kluwer Academic Publishers, Dordrecht.

King, D.A. and J.A. Sinden (1988) "Influence of Soil Conservation on Farm Land Values", *Ecological Economics*, 64(3), pp. 242-255.

Lacy, R.C. (1993) "VORTEX: a computer simulation model for population viability analysis", *Wildlife Research,* 20, pp. 45-65.

Lacy, R.C and T.W. Clark (1990) "Population viability assessment of eastern barred bandicoot, in T.W. Clark and J.H. Seebeck (eds) *Management and Conservation of Small Populations,* Chicago Zoological Society, Brookfield, IL.

Levin, S.A. (1998) "Ecosystems and the Biosphere as Complex Adaptive Systems", *Ecosystems*, 1(5), pp. 431-436.

Levin, S.A., Greenfell, B., Hastings, A. and A. Perelson (1997) "Mathematical and Computational Challenges in Population Biology and Ecosystem Science", *Science*, 275, pp. 334-343.

Lindenmayer, D.B., Lacy, R.C., Thomas, V.C. and T.W. Clark (1993) "Predictions of the impacts of changes in population size and environmental variability on Leadbeater's possum using population viability analysis", *Wildlife Research*, 20, pp. 67-86.

Lockwood, M. (1999) "Humans Valuing Nature: Synthesising Insights from Philosophy, Psychology and *Environment Values*, 8, pp. 381-401.

Loomis, J.B. and D.M. Larson (1994) "Total economic value of increasing Gray Whale populations: results from a contingent valuation survey of visitors and households", *Marine Resource Economics*, 9, pp. 275-286.

Mace, G. M. et al. (1992) "The development of new criteria for listing species on the IUCN Red List" *Species*, 19, pp. 16-22.

Mace, G. M., and R. Lande. (1991) "Assessing extinction threats: toward a reevaluation of IUCN threatened species categories", *Conservation Biology*, 5(2), pp. 148-57.

Mace, G. M. & S. N. Stuart. (1994) "Draft IUCN Red List Categories, Version 2.2", *Species*, 21-22, pp. 13-24.

Maynard-Smith, J. (1974) *Models in Ecology.* Cambridge University Press, Cambridge.

Moons, E. (1999) "Estimation of the recreation values of a forest", *M.Sc. dissertation,* Faculty of Economics and Applied Economics, Catholic University of Leuven.

NOAA - National Oceanic and Atmospheric Administration (1993) "Report of the NOAA Panel on Contingent *Federal Register*, 58(10), pp. 4601-4614, US.

Norberg, J. (1999) "Linking Nature´s Services to Ecosystems: Some General Ecological Concepts", *Ecological Economics*, 29, pp. 183-202.

Nordhaus, W.D. (1991) To slow or not to slow: The economics of the greenhouse effect. *Economic Journal*, Vol. 101: 920-937.

Nordhaus, W.D. (1994) *Managing the Global Commons: The Economics of Climate Change*. The MIT Press, Cambridge.

Norton, B. G. (1982) "Environmental Ethics and Nonhuman Rights", *Environmental Ethics*, 4, 17-36.

Norton, G. M. and C. Southey (1995) "The opportunity costs of biodiversity protection in Kenya", *Ecological Economics,* 11, pp. 135-151.

, Chapter IV, Elsevier Science Publishers, North-Holland.
Parry, M., and T. Carter (1998) *Climate impact and adaptation assessment*. Earthscan, London.
Patten, B.C. (ed.) (1971) *Systems Analysis and Simulation in Ecology*, Volumes I-IV, Academic Press.
Pearce, D. (1999) "Valuing biological diversity: issues and overview", draft paper presented at *OECD Workshop on Benefit Valuation for Biodiversity Resources*, Paris, France.
Pearce, D. and D. Moran (1994) "The Economic Value of Biodiversity", Earthscan Publications Limited, London, UK.
Perrings, C. (1998) "Resilience in the dynamics of economic-environment systems", *Environmental and Resource Economics: Special Issue*, 11, pp. 503-520.
Perrings, C., K-G. Maler, C. Folke, C.S. Holling and B-O. Jansson (eds) (1995) "Biodiversity Loss: Economic and Ecological Issues", Cambridge University Press.
Perrings, C. and B. Walker (1997) "Biodiversity, resilience and the control of ecological-economic systems: the case of fire-driven rangelands", *Ecological Economics* 22, pp. 73-83.
Randwell, D.S. (1969) pers. comm. cited in Ratcliffe, D.A. (1986) "Selection of important areas for wildlife conservation in Great Britain: the Nature Conservancy Council's Approach, in M.B. Usher (ed.), *Wildlife Conservation Evaluation*, Chapman & Hall, London, UK.
Ribaudo, M.O. (1989) "Targeting the Conservation Reserve Program to Maximize Water Quality Benefits", *Land Economics*, 65(4), pp. 320-332.
Richer, J. (1995) " Willingness to Pay for Desert Protection", *Contemporary Economic Policy*, XIII, pp. 93-104.
RIZA (1999) "Uitganspunten voor een Graadmeteer Natuur-waarde voor Nederland: het Ecologisch Kapitaal Index concept, definitiestudie in opdracht van het Natuurplanbureau", unpublished.
Rotmans, J., and B. de Vries (1997) *Perspectives on Global Change: The Targets Approach*. Cambridge University Press, Cambridge, UK.
Roughgarden, R., R.M. May and S.A. Levin (eds.), 1989. *Perspectives in Ecological Theory*. Princeton University Press, Princeton, NJ.
Ruijgrok, E. (1999) "Valuation of Nature in Coastal Zones", *Ph.D. dissertation*, Free University of Amsterdam.
Sagoff, M. (1980) "On the Preservation of Species", *Columbia Journal of Environmental Law*, 7, pp. 33-67.
Samples, K. and J. Hollyer (1989) "Contingent Valuation of Wildlife Resources in the Presence of Substitutes and Complement", in Johnson, R. and G. Johnson (eds) "Economic Valuation of Natural Resources: Issues,

Shackley, S. (1997) "Trust in models? The mediating and transformative role of computer models in environmental discourse", in: M. Redclift and G. Woodgate (eds.). *The International Handbook of Environmental Sociology*, Edward Elgar Publ., Cheltenham.
Sijtsma, F.J.and D. Strijker (1995) "Effect analyse ecologische hoodstructuur; Deel II: Natuurwaarde, Stiching Ruimtelijke Economie Groningen, Groningen, The Netherlands.
Sijtsma, F.J., Strijker, D. and M.L. Hoefsloot (1998) "Duurzame Ontwikkeling in het Waddengebied", University of Groningen, Groningen, The Netherlands.
Silberman, J., Gerlowski, D.A. and N.A. Williams (1992) "Estimating Existence Value for Users and Nonusers *Land Economics*, 68(2), pp. 225-236.
Soulé, M.E. (ed) (1987) *Viable Population for Conservation*, Cambridge University Press, Cambridge, UK.
Spellerberg, I.F. (1992) *Evaluation and Assessment for Conservation*, Chapman & Hall, London, UK.
Swallow, S.K. (1994) "Renewable and nonrenewable resource theory applied to coastal agriculture, forest, wetland and fishery linkages", *Marine Resource Economics* 9(4), pp. 291-310.

Tol, R. (1998) Economic aspects of global environmental models. In: J.C.J.M. van den Bergh and M.W. Hofkes (eds.). *Theory and Implementation of Economic models for Sustainable Development*. Kluwer Academic Publishers, Dordrecht.

Turner, R.K., Folke, C., Gren I.M. and I.J. Batemann (1995) "Wetland Valuation: Three Case Studies", in Perrings C., Mäler, K-G., Folke, C., Holling, C.S. and B-O. Jansson (eds), *Biodiversity Loss: Economic and Ecological Issues*, Cambridge University Press.

Turner, R.K., Perrings, C. and Folke, C. (1997) *Ecological Economics: Paradigm or Perspective?* in J. Van den Bergh and J. Van der Straaten (eds.) Economy and Ecosystems in Change, Edward Elgar Publishers, Cheltenham.

Turner, R.K., K. Button and P. Nijkamp (eds) (1999) "Ecosystems and Nature: Economics, Science and Policy", Elgar.

Ulanowicz, R.E. (1992) "Ecosystem Health and Tropic Flows Networks" in Costanza, R., Norton, B. and B. Haskell (eds), *Ecosystem Health: New Goals for Environmental Management*, Island Press, Washington, US.

UNEP (1995) "Global Biodiversity Assessment" report by Watson, R.T., Heywood, V.H., Baste, I., Dias, B., Gamez, R. and W. Reid, United Nations Environment Program, US.

UNEP (1997) "Recommendations for a Core Set of Indicators of Biological Diversity", Convention on Biological Diversity, Montreal, United Nations Environment Program, US.

Usher, M.B. (1989) "Nature Conservation in UK", *Journal of Applied Ecology*, 26, pp. 813-824.

van den Bergh (1996) *Ecological Economics and Sustainable Development: Theory, Methods and Applications.* Edward Elgar, Cheltenham, UK.

van den Bergh (2000) "Integrated modelling and assessment for ecosystem management", *mimeo*, Department of Spatial Economics, Amsterdam Free University.

van den Bergh and P. Nijkamp (1991) "Aggregate economic-ecological models for sustainable development", *Environment and Planning*, 23: pp. 187-206.

van den Bergh and P. Nijkamp (1994) "An Integrated Model for Economic Development and Natural Environment: An Application to the Greek Sporades Islands", *The Annals of Operations Research* 54: 143-174.

van den Bergh, J.C.J.M. van den, and M.W. Hofkes (eds.), (1998) *Theory and Implementation of Economic Models for Sustainable Development.* Kluwer Academic Publishers, Dordrecht.

van den Bergh and J. van der Straaten (eds) (1997) Economy and Ecosystems in Change, Edward Elgar Publishers, Cheltenham.

van der Veer D. and C. Pierce (1986) "People, Penguins and Plastic Trees: Basic Issues in Environmental Wadsworth, Belmont, US.

van Kooten, G.C. (1993) " Bioeconomic Evaluation of Government Agricultural Programs on Wetland *Land Economics*, 69(1), pp. 27-38.

Watt, K.E.F. (1968) *Ecology and Resource Management: A Quantitative Approach.* McGraw-Hill, New York.

Whittaker, R.H. (1960) "Vegetation of the Siskiyou Mountains, Oregon and California", *Ecological Monographs*, 30, pp. 279-338.

Whittaker, R.H. (1972) "Evolution and Measurement of Species Diversity", *Taxon*, 21, pp. 1-67.

Williams, P.H., Vane-Wright, R.I. and RI Humphries (1991) "Measuring Biodiversity for Choosing Conservation Areas" in LaSalle (ed) *Hymenoptera and Biodiversity*, CAB International.

WTO (1997) *Statistical Year Book of Tourism*, World Tourism Organisation, Madrid.

Zhou, Z. and W. Pan (1997) "Analysis of the viability if a giant panda population", *Journal of Applied Ecology*, 34, pp. 363-374.

Zuchetto, J. and A.M. Jansson (1985) *Resources and Society: A Systems Ecology Study of the Island of Gotland, Sweden.* Springer-Verlag, New York.

Zurlini, G., Amadio, G. and O. Rossi (2000) "A Landscape Approach to Biodiversity and Biological Planning: the Map of Italian Nature, *Ecosystem Health*, 5(4), pp. 294-311.

Right to Property, Investments, and Environmental Protection: The Perspectives of the European and Inter-American Courts of Human Rights*

Cesare Pitea

1. INTRODUCTORY REMARKS

The tension between the protection of foreign investments and the protection of the environment[1] is often perceived as a conflict between investors' rights and the 'right to regulate' (or even the 'obligation to regulate' under international law) of the host state. However, neither international investment law prevents states from taking property for environmental protection purposes, nor it places a duty upon them to refrain from enacting regulatory measures to ensure that economic activities, including foreign investments, are compatible with their desired standard of environmental protection. It only requires respect for certain conditions, including, as the case may be, the payment of compensation. Therefore, the real matter is to identify in which cases the exercise of sovereign rights to regulate affecting foreign investments carries with it – together with other conditions – a duty to pay compensation, and the nature and amount of the latter.[2] This issue is better conceptualized as one relating to the allocation of costs of environmental policies, rather than a conflict between state sovereignty and individual rights of investors. In other words, it is a matter of distributive justice.

Looking at similar problems from two very different (if not opposite) angles, the European and American regional courts of human rights, have developed some common jurisprudential patterns. After describing the approaches of the two Courts, I will single out three areas of convergences and the relevant principles deriving from the case-law,[3]

* The present paper is a draft of the Chapter to be published in T. TREVES, F. SEATZU, S. TREVISANUT (eds), Foreign Investment, International Law and Common Concerns, Routledge, Oxford, 2013, pp. 265-279.

[1] See, generally, R. Pavoni, 'Environmental Rights, Sustainable Development, and Investor-State Case-Law: A Critical Appraisal', in P.-M. Dupuy et al. (eds.) Human Rights in International Investment Law and Arbitration, Oxford: OUP, 2009, pp. 525-56; and J. E. Viñuales, 'Foreign Investment and the Environment in International Law: An Ambiguous Relationship', British Yearbook of International Law, 2009, vol. 80, pp. 244-332.

[2] On the sovereign rights to regulate, see in this volume the chapter by P. Bertoli and Z. Crespi Reghizzi, and A. Viterbo. Specifically on environmental expropriation, see in this volume the chapter by P. Vargiu.

[3] In this sense, the present chapter does not aim at comparing systematically the jurisprudence of human rights courts and international investment arbitral tribunals. For an analysis of the potential and limits for the protection of foreign investments in the European system of human rights see, M. Ruffert, 'The Protection of Foreign Direct Investment by the European Convention on Human Rights', German Yearbook of International Law, 2000, vol. 43, pp. 116-48; and P. De Sena, 'Economic and Non-Economic Values in the Case-Law of the European Court of Human Rights', in Dupuy et al. (eds), op. cit., pp. 208-18. For an analysis centred on the judicial protection of foreign investments through the European Court of Human Rights, see C. Tomuschat, 'The European Court of Human Rights and Investment Protection', in C. Binder et al. (eds) International Investment Law for the 21st Century: Essays in Honour of Christoph Schreuer, New York: OUP, 2009, pp. 637-56 and U. Kriebaum, 'Is the European Court of Human Rights an Alternative to Investor-State Arbitration?', in Dupuy et al. (eds), op. cit., pp. 219-45.

DRAFT

with the aim of stimulating further reflection by 'international investment lawyers' on the possibility of drawing inspiration from this jurisprudence.

2. TWO PERSPECTIVES ON THE RELATIONSHIP BETWEEN THE RIGHT TO PROPERTY, INVESTMENTS, AND ENVIRONMENTAL PROTECTION

For the sake of the argument, and to a some extent artificially, I will describe the European Court of Human Rights (ECtHR) as the champion of a 'traditional' perspective on the relationship between the right to property and environmental concerns, and the Inter-American Court of Human Rights (IACtHR) as the forum where a 'reverse' perspective has been taken.[4]

According to the 'traditional' perspective environmental measures, whether regulatory or administrative in nature, may amount to interferences in 'possessions' protected by the right to property, a notion that may overlap with that of 'investment'.[5] As it is widely known, in order to be compatible with Article 1 of Protocol No. 1 to the European Convention on Human Rights (ECHR), these interferences must satisfy three conditions: legality (they must be in accordance with the law); legitimate aim (they must pursue a general interest of the society at large); and proportionality (domestic authorities, within their margin of appreciation, must struck a fair balance between the competing public and private interest, without impinging on the very essence of the right of the individual concerned).[6] The 'common concern' to environmental protection constitutes a limitation to investors' right to exploit their investment and, in some cases, to claim compensation for the reduction of its profitability.

The 'reverse' perspective looks at investment plans and projects themselves as interferences with property rights, after having redefined the latter beyond the traditional boundaries of an individual right based on a legal title, to encompass a collective dimension. This notion emerged in the context of the international law on the protection of indigenous and tribal peoples,[7] in order to protect them against economic

DRAFT

[4] In fact, the Inter-American Court of Human Rights (IACtHR) has also applied the right to property in order to protect foreign investors' rights, although this role is limited by the lack of locus standi of legal persons before the organs of the Inter-American system; see P. Nikken, 'Balancing of Human Rights and Investment Law in the Inter-American System of Human Rights', in Dupuy et al. (eds), op. cit., pp. 255-259.

[5] See, C. H. Schreuer and U. Kriebaum, 'The Concept of Property in Human Rights Law and International Investment Law', in S. Breitenmoser et al. (eds) Human Rights, Democracy and the Rule of Law: Liber Amicorum Luzius Wildhaber, Baden-Baden: Nomos, 2007, pp. 743-62.

[6] See, recently, Depalle v. France [GC], no. 34044/02, Judgment of 29 March 2010 (Merits and Just Satisfaction), para. 83; Brosset-Triboulet and Others v. France [GC], no. 34078/02, Judgment of 29 March 2010 (Merits and Just Satisfaction), para. 86.

[7] ILO Convention No. 169 Concerning Indigenous and Tribal Peoples in Independent Countries (27 June 1989), entered into force on 5 September 1991, 72 ILO Official Bull. 59 [hereinafter ILO Convention No. 169]; United Nations Declaration on the Rights of Indigenous Peoples (New York, 13 September 2007), U.N. Doc. A/RES/61/295 (2 October 2007) [hereinafter UN DRIP]. See, generally, S. J. Anaya, Indigenous Peoples and International Law, 2nd ed., New York: OUP, 2004, pp. 141-8. For an analysis of the issue, see the working paper by the Special Rapporteur to the UN Sub-Commission on the Promotion

uses of their ancestral lands and natural resources by third parties that may be prejudicial to their physical and cultural survival. This approach has found judicial protection and development[8] in the inter-American system of human rights, on the basis of Article 21 of the American Convention on Human Rights (ACHR).[9] A growing case-law and a consistent jurisprudence of the IACtHR[10] recognize proprietary rights to indigenous and tribal communities over ancestral territories and natural resources of traditional use located therein,[11] even in absence of a formal title, on the basis of traditional possession.[12] This line of reasoning assumes a relevant environmental dimension by virtue of the particular relationship between these cultures and nature.[13] Through the lens of the protection of indigenous peoples, the right to property is used as a tool for protecting the 'common concern' for environmental protection, against the interests (and rights) of actual or prospective owners and investors.

In both perspectives, the focus is on how to strike the balance between the competing interests at stake. In doing so the two Courts have taken some stances that, notwithstanding differences, point at the same direction in giving relevance to certain elements. The need to introduce a balancing test in the reasoning of investment tribunals, mainly in the form of a proportionality test, is increasingly called for by

---

and Protection of Human Rights, Ms. Erica-Irene Daes, on Indigenous Peoples and Their Relationship to Land, UN Doc. E/CN.4/Sub.2/2001/21 (2001).

[8] See G. Pentassuglia, 'Towards a Jurisprudential Articulation of Indigenous Claim Rights', European Journal of International Law, 2011, vol. 22, pp. 165-202, pp. 177-8 (arguing that the standard resulting from the Inter-American jurisprudence goes further than the one available under the ILO Convention no. 169 and the UN DRIP in several respects). For an analysis pointing at shortcomings in this jurisprudence, see J. M. Pasqualucci, 'International Indigenous Land Rights: A Critique of the Jurisprudence of the Inter-American Court of Human Rights in Light of the United Nations Declaration on the Rights of Indigenous Peoples', Wisconsin International Law Journal, 2009, vol. 27, pp. 51-98.

[9] '[A]rticle 21 of the Convention protects the right to property in a sense which includes, among others, the rights of members of the indigenous communities within the framework of communal property', Mayagna (Sumo) Awas Tingni Community v. Nicaragua, Judgment of 31 August 2001 (Merits, Reparations and Costs), para. 148.

[10] Mayagna (Sumo) Awas Tingni Community v. Nicaragua, cit.; Moiwana Community v. Suriname, cit.; Yakye Axa Indigenous Community v. Paraguay, Judgment of 17 June 2005 (Merits, Reparations and Costs); Sawhoyamaxa Indigenous Community v. Paraguay, Judgment of 29 March 2006 (Merits, Reparations and Costs); Saramaka People v. Suriname, Judgment of 28 November 2007 (Preliminary Objections, Merits, Reparations and Costs), and Judgment of 12 August 2008 (Interpretation of the Judgment of Preliminary Objections, Merits, Reparations and Costs); Xákmok Kásek Indigenous Community v. Paraguay, Judgment of 24 August 2010 (Merits, Reparations and Costs); and Kichwa Indigenous People of Sarayaku v. Ecuador, Judgment of 27 June 2012 (Merits and Reparations). This chapter will not deal, given the limited space available, with the practice of the Inter-American Commission on Human Rights, although the fundamental contribution of this organ in shaping the concepts and standards relating to the protection of indigenous and tribal peoples in the inter-American system should not be overlooked. For a complete overview, see Inter-American Commission on Human Rights, 'Indigenous and Tribal Peoples' Rights Norms and Jurisprudence of the Inter-American Human Rights System', OEA/Ser.L/V/II., Doc. 56/09, 2009.

[11] Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 137; Saramaka People v. Suriname (Merits), op. cit., para. 121.

[12] Mayagna (Sumo) Awas Tingni Community v. Nicaragua, op. cit.

[13] On the environmental aspects of indigenous rights, see A. Fodella, 'Indigenous Peoples, the Environment, and International Jurisprudence, in N. Boschiero et al. (eds.) International Courts and the Development of International Law. Essays in Honour of Tullio Treves, The Hague: T.M.C. Asser Press, 2013, pp. 349-364.

DRAFT

scholars[14] and to some extent emerging from arbitral practice. In this perspective, the common elements taken into account by the two human rights courts in doing the same exercise may be a useful source of inspiration for international investment lawyers.

3. HIERARCHY OF UNDERLYING VALUES

The first common point is the inherently different weight attached to the interests at stake in the judicial reasoning of the two courts.

In the Sawhoyamaxa Community v. Paraguay case, the IACtHR has made an attempt at approaching this issue within the conceptual framework of a formal hierarchy of sources of international law, pointing at the superior rank of the ACHR, as a multilateral and normative treaty for the protection of human rights, over bilateral investment treaties (BITs). The case concerned the reclamation by an indigenous community of traditional lands allegedly owned under Paraguayan law by German nationals and thus protected under a BIT between the two states. The IACtHR has overtly stated that the enforcement of such agreements (that she characterizes as 'commercial' legal instruments)

> 'should always be compatible with the American Convention [on Human Rights], which is a multilateral treaty on human rights that stands in a class of its own and that generates rights for individual human beings and does not depend entirely on reciprocity among States.'[15]

This reasoning, reflecting the idea that international human rights law is inherently superior to other branches of international law,[16] is rather weak. It is difficult to deny that treaty commitments entered into by States have under international law the same formal status and rank regardless of their subject-matter.[17] Moreover, even considering that this argument may point at a possible development of international law, it is very unlikely to find any international tribunal other than a human rights court willing to follow this line of reasoning.

It is by no means surprising that the ICSID Arbitral Tribunal in the Suez v. Argentina case rejected a similar argument suggesting that:

---

[14] See, among others, B. Kingsbury and S. W. Schill, 'Public Law Concepts to Balance Investors' Rights with State Regulatory Actions in the Public Interest – the Concept of Proportionality', in S. W. Schill (ed.) International Investment Law and Comparative Public Law, New York: OUP, 2010, pp. 75-104; J. Krommendijk and J. Morijn, '"Proportional" by what Measure(s)? Balancing Investor Interests and Human Rights by Way of Applying the Proportionality Principle in Investor-State Arbitration in Human Rights in International Investment Law and Arbitration', in Dupuy et al. (eds), op. cit., pp. 42-51; A. Tanzi, 'On Balancing Foreign Investment Interests with Public Interests in Recent Arbitration Case-Law in the Public Utilities Sector', The Law and Practice of International Courts and Tribunals, 2012, vol. 11, pp. 47-76.

[15] Sawhoyamaxa Indigenous Community v. Paraguay, op. cit., para. 140.

[16] According to one scholar, '[t]his activist approach permeates much of the Court's activities, and it proposes some kind of hierarchy within international law in which human rights law is at the top', see L. Lixinski, 'Treaty Interpretation by the Inter-American Court of Human Rights: Expansionism at the Service of the Unity of International Law', European Journal of International Law, 2010, vol. 21, p. 590.

[17] See, Nikken, op. cit., pp. 265-70.

DRAFT

> 'Argentina's human rights obligations to assure its population the right to water somehow trumps its obligations under the BITs and that the existence of the human right to water also implicitly gives Argentina the authority to take actions in disregard of its BIT obligations. The Tribunal does not find a basis for such a conclusion either in the BITs or international law. Argentina is subject to both international obligations, i.e. human rights and treaty obligation, and must respect both of them equally.'[18]

Of greater interest is the role played by value judgments and considerations in defining the standard, the limits and the process of the balancing exercise.

Although a 'human right to an environment of quality' is neither affirmed in the ECHR and its Protocol, nor is clearly established as a matter of customary international law, the ECtHR places a great emphasis on the fundamental nature of the common interest to environmental protection. In the case of Hamer v. Belgium, the ECtHR has expressed its views in the clearest terms: 'Financial imperatives and even certain fundamental rights, such as ownership, should not be afforded priority over environmental protection considerations'.[19] This approach is further reflected in several steps of the Court's reasoning in the relevant cases, based on the assumption that 'environmental conservation policies, where the community's general interest is pre-eminent, confer on the State a margin of appreciation that is greater than when exclusively civil rights are at stake.'[20]

When dealing with large development projects relating to, or impinging on, natural resources located within indigenous communities' traditional lands, the IACtHR has stressed the pre-eminence of the interest at stake for the peoples concerned. In view of their special relationship with the natural environment, the guarantee and protection of property rights of indigenous and tribal peoples constitutes a precondition for their survival. The meaning of this term goes beyond the mere physical existence of those peoples and encompasses 'their distinct cultural identity, social structure, economic system, custom, beliefs and traditions'.[21] Therefore, in the interpretation judgment in the Saramaka People v. Suriname case, the Court found that:

> 'the term "survival" (…) does not refer only to the obligation of the State to ensure the right to life of the victims, but rather to (…) the continuance of the relationship of the [indigenous peoples] with their land or their culture.'[22]

In this context, the relevance of protecting the ancestral territories of these communities is reinforced by the fact that the protection of proprietary rights is inextricably linked to the guarantee of other rights, of an absolute or even cogent nature, such as the right to life (Article 4 ACHR) and the protection from torture and inhuman treatment (article 5

[18] Suez, Sociedad General de Aguas de Barcelona, S.A. and Vivendi Universal, S.A. v. Argentine Republic, ICSID Case no. ARB/03/19, Decision on Liability of 30 July 2010, para. 262. For a comment on international investment arbitrations in the water services sector, see in this volume the chapter by A. Tanzi.

[19] Hamer v. Belgium, no. 21861/03, Judgment of 27 November 2007 (Merits and Just Satisfaction), para. 75. See, also, Turgut and Others v. Turkey, no. 1411/03, Judgment of 8 July 2008 (Merits), para. 90.

[20] Depalle v. France, op. cit., para. 84; Brosset-Triboulet and Others v. France, op. cit., para. 94.

[21] Saramaka People v. Suriname (interpretation), para. 37.

[22] Ibid., para. 29.

DRAFT

ACHR).[23] In fact, separation from ancestral lands bears significantly on access to basic resources and means of subsistence, such as water and food,[24] and on important spiritual elements of indigenous cultures, such as the possibility to perform traditional funerary rituals.[25] In those cases, the above-mentioned (and other) rights may be affected and violated.

Therefore, in deciding cases relating to indigenous and tribal peoples the IACtHR takes into account the 'special meaning of communal property of ancestral lands for the indigenous peoples, including the preservation of their cultural identity and its transmission to future generations relationship'[26] and compatibility of those interferences with indigenous and tribal peoples 'survival' – as broadly defined – is 'a crucial factor' in assessing the permissibility of an interference in their territorial rights.[27]

4. FLEXIBILITY IN CONSIDERING THE ROLE OF COMPENSATION FOR INTERFERENCES WITH PROPERTY RIGHTS

The role played by compensation in assessing compliance of the conduct of states with their international human rights obligations is strictly linked with the normative consideration of the inherent difference in weight of the interests concerned.

In the jurisprudence of the ECtHR the amount of compensation paid domestically is a key factor in assessing the proportionality of a lawful[28] interference with the right to property, requiring that an excessive burden is not placed on the person concerned. However, its relevance varies greatly depending on whether the interference amounts to deprivation or regulation of property. In the first case, the ECtHR has recently summarized its position as follows:

---

[23] See, K. Rinaldi, 'Le droit des populations autochtones et tribales à la propriété dans le système interaméricain de protection des droits de l'homme', in L. Hennebel and H. Tigroudja (eds.) Le *particularisme interaméricain des droits de l'homme : en l'honneur du 40e anniversaire de la Convention américaine des droits de l'homme*, Pedone: Paris, 2009, pp. 244-7.

[24] See, especially, Sawhoyamaxa Indigenous Community v. Paraguay, op. cit. The IACtHR has interpreted the right to life as encompassing a social dimension, since it comprises a 'the right that conditions that impede or obstruct access to a decent existence should not be generated' (Yakye Axa Indigenous Community v. Paraguay, op. cit., paras 161-162). Therefore the Court has directly linked territorial deprivation with the protection of the rights to life, founding that 'denying [indigenous communities] their sources of nutrition and breaking their symbolistic relationship with their land, has a deleterious effect on their health' (ibid., para. 163). However, in the latter case the Court was unable to establish a causal link between such life conditions and the death of several community members. This finding, largely criticized by the then President of the Court, Antonio Cançado Trindade in his dissenting opinion, has been partially redressed in the subsequent Sawhoyamaxa judgment.

[25] See, Moiwana Community v. Suriname, op. cit., paras 90-103.

[26] Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 124.

[27] Saramaka People v. Suriname (merits), op. cit., para. 128. The Human Rights Committee has taken a similar position on the compatibility of development measures with the rights of minorities under art. 27 of the International Covenant on Civil and Political Rights, see Länsman et al. v. Finland, Comm. No. 511/1992, UN Doc. CCPR/C/52/D/511/1992 (1994), para. 9.4.

[28] Unlawful interferences, whether amounting to expropriation or not, determine a breach of the ECHR, calling for restitutio in integrum in accordance with general international law on State responsibility, see Guiso-Gallisay v. Italy [GC], no. 58858/00, Judgment of 22 December 2009 (Just Satisfaction).

DRAFT

> 'the taking of property without payment of an amount reasonably related to its value will normally constitute a disproportionate interference. Article 1 of Protocol No. 1, however, does not guarantee a right to full compensation in all circumstances. Legitimate objectives of "public interest" may call for less than reimbursement of the full market value of the expropriated property.'[29]

While the pursuance of a legitimate public interest is always required for the deprivation of property to comply with the conventional standard, only in some cases less than full compensation may be held to maintain the required balance. This applies, for instance, when restricting measures are taken in the context of broad social and economic reforms[30], and of significant transition or change in political and constitutional regimes.[31] The question is whether other situations may be invoked. In the *Kozacıoğlu* v. Turkey judgment – from which the citation above is taken – the Grand Chamber of the Court has found that other situations may be relevant, taking into account the inherent nature of the general interest at stake. In fact, it has expressed the view that 'the protection of the historical and cultural heritage is one such objective [justifying less than reimbursement of the full market value of the expropriated property]', and environmental protection would possibly enjoy the same status.[32] This position sharply contrast with the principle affirmed, by the ICSID Arbitral Tribunal in the Santa Elena v. Costa Rica award, according to which:

> 'Expropriatory environmental measures – no matter how laudable and beneficial to society as a whole – are, in this respect, similar to any other expropriatory measures that a state may take in order to implement its policies: where property is expropriated, even for environmental purposes, whether domestic or international, the state's obligation to pay compensation remains. (…) there is no dispute between the parties as to the applicability of the principle of full compensation for the fair market value of the Property (…)'.[33]

Moreover, one should not lose sight of the proviso made in Article 1 of Protocol No. 1, to the effect that deprivation of property is 'subject to the conditions provided for (…) by the general principles of international law'. The Court has interpreted this provision as applying solely to interferences with possessions of non-nationals of the state concerned and it has justified the higher standard of protection of aliens on the basis of a supposed vulnerability of foreigners to domestic legislation, since they do not contribute through political participation to the shaping of national policies, especially in the field of socio-economic reforms.[34] This position, expressed by way of obiter dicta, has never been tested in actual cases of expropriation of foreigners' possessions. However, it is reasonable to consider that it would be maintained in relevant future

DRAFT

[29] *Kozacıoğlu* v. Turkey [GC], no. 2334/03, Judgment of 19 February 2009 (Merits and Just Satisfaction), para. 64.
[30] Lithgow and Others v. United Kingdom, nos. 9006/80 et al., Judgment of 8 July 1986 (Merits).
[31] See Broniowski v. Poland [GC], no. 31443/96, Judgment of 22 June 2004 (Merits), para. 182; The Former King of Greece and Others v. Greece [GC], no. 25701/94, Judgment of 23 November 2000 (Merits), para. 88; Forrer-Niedenthal v. Germany, no. 47316/99, Judgment of 20 February 2003 (Merits), para. 48.
[32] *Kozacıoğlu* v. Turkey, op. cit..
[33] Compañiá del Desarrollo de Santa Elena, S.A. v. The Republic of Costa Rica, ICSID Case No. ARB/96/1, Final Award of 17 February 2000, paras 71-72.
[34] James and Others v. United Kingdom, Judgment of 21 February 1986 (Merits), para. 63 and Lithgow and Others v. United Kingdom, op. cit., para. 116.

cases, so that this class of applicants may claim to be entitled to full compensation in any case.[35]

With regard to interferences not amounting to deprivation of property, the ECtHR has taken the view that 'the lack of compensation is a factor to be taken into consideration in determining whether a fair balance has been achieved but is not of itself sufficient to constitute a violation of Article 1 of Protocol No. 1'.[36] Therefore, in this context, a case-by-case assessment of the relevance of the domestic attitude towards compensation and determination thereof shall be made and states appear to enjoy a broader margin of appreciation. Moreover, foreigners cannot take advantage of any higher standard of protection possibly available under general international law, since the above-mentioned provision applies only to deprivation of property and not to other forms of interference with the protected right.

The weight of this general consideration increases if one looks specifically at the criteria used by the ECtHR to distinguish between taking and regulation, which are different from those used in the context of investment protection and may lead to different results. A recent example is particularly significant. In the Consorts Richet and Le Ber v. France case, the applicants had sold to the state their properties on the Porquerolles island, at a price below the market value, following a contractual undertaking by the state to permit a certain amount of construction over the portions of land retained by the applicants. Some twenty years later, before the realization of most of the projected development, a construction ban for environmental conservation reasons was imposed on the region comprising the island. The Court characterized the measure as one of control on the use of property, notwithstanding the fact that rights of the applicants, based on previous specific commitments contractually undertaken by the State, had been rendered void in practice.[37] In the same conditions, had the applicants been foreign nationals, the situation could have been considered an indirect expropriation under international investment law, calling for prompt, adequate and effective compensation.[38]

Not only this case indicates a potential divergence between the protection of property under the ECHR and the law on international investments. More importantly, it demonstrates that recourse to a flexible approach allows the ECtHR to take into account all the circumstances of each case, including the allocation of the costs for environmental protection policies, in order to determine whether a violation of the right

[35] See U. Kriebaum, 'Is the European Court of Human Rights an Alternative to Investor-State Arbitration?', in Dupuy et al. (eds), op. cit., pp. 242-3.
[36] Brosset-Triboulet and Others v. France, op. cit., para. 94.
[37] Consorts Richet and Le Ber v. France, nos. 18990/07 and 23905/07, Judgment of 18 November 2010 (Merits and Just Satisfaction), para. 113.
[38] See, for example, in the context of the North American Free-Trade Agreement (NAFTA): Metalclad Corporation v. The United Mexican States, ICSID Case No. ARB(AF)/97/1, Award of 30 August 2000, para. 103 ('expropriation under NAFTA includes (…) also covert or incidental interference with the use of property which has the effect of depriving the owner, in whole or in significant part, of the use or reasonably-to-be-expected economic benefit of property'); and, a contrario, Methanex Corporation v. United States of America (UNCITRAL), Final Award on Jurisdiction and Merits of 3 August 2005, Part IV, Chapter D, p. 4, para. 4 ('as a matter of general international law, a non-discriminatory regulation for a public purpose (…) is not deemed expropriatory and compensable unless specific commitments had been given by the regulating government to the then putative foreign investor contemplating investment that the government would refrain from such regulation').

DRAFT

of property has occurred. In fact, the ECtHR finally observed that, had the ban on the specific constructions to which the applicants were entitled been necessary to preserve the environment, a fair balance between the competing interest could have been ensured by recognizing to the applicants 'une compensation matérielle ou financière en réparation du préjudice subi du fait du non-respect des actes de vente'.[39] Failing compensation, the applicants' rights had been violated and a (quite substantial) sum was awarded as just satisfaction by the ECtHR.

Under Article 21 of the ACHR, expropriation of, and restriction to, property are generally allowed under conditions that are similar to those applicable in the European context, including the payment of compensation.[40] However, compensation for interferences with the right to property is framed differently when ancestral territories of indigenous communities are at stake, given the peculiar value and functions of the continued relationship of those communities with their land.

When public authorities promote or authorize a project concerning the exploitation of natural resources within traditional indigenous lands, compliance with the standard of protection deriving from the right to property requires additional conditions to be met. Firstly, such projects interfering with traditional territorial rights must not amount to a 'denial [of indigenous peoples'] traditions and customs in a way that endangers the very survival of the group and of its members'.[41] Along with this substantive standard, the IACtHR also requires that indigenous communities are consulted, that – in certain cases – they give their free, prior, and informed consent to the project, that an equitable share of the project's benefits is ensured to the community, and that an environmental and social assessment is undertaken before authorizing the exploitation.[42] Unlike the case of ordinary interferences through lawful measures taken in the public interest, full monetary compensation is not as such sufficient to bring a limitation of indigenous peoples' property rights in line with the conventional standard.

When the peoples concerned have unwillingly lost possession of their traditional lands, the Court has consistently held that the state must guarantee through domestic legislation, as well as administrative and judicial arrangements, the legal restitution of the land concerned and its physical return to the interested people. As it has been observed:

> 'Due to its nature, the indigenous peoples' property is a non-fungible asset, with regard to its economic value, so the payment of compensation in the event of an expropriation does not repair in any way the damage caused to the cultural assets that are part of the property.'[43]

The IACtHR has also elaborated some criteria on how to deal with restitution 'when indigenous communal property and individual private property [of third parties] are in real or apparent contradiction',[44] in particular when 'the lands have been lawfully

[39] Ibid., para. 122.
[40] Ivcher-Bronstein v. Peru, Judgment of 6 February 2001 (Merits, Reparations and Costs), para. 128.
[41] Saramaka People v. Suriname (merits), op. cit., par. 128
[42] Ibid., paras 128 and 143.
[43] Nikken, op. cit., p. 262.
[44] Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 144.

DRAFT

transferred to third parties in good faith'.[45] Although the IACtHR admits to lack the competence to decide upon these 'domestic' matters,[46] she has observed that:

> 'the fact that the claimed lands are privately held by third parties is not in itself an 'objective and reasoned' ground for dismissing prima facie the claims by the Indigenous people. Otherwise, restitution rights become meaningless and would not entail an actual possibility of recovering traditional lands, as it would be exclusively limited to an expectation on the will of the current holders, forcing indigenous communities to accept alternative lands or economic compensations'.[47]

Ordinary principles on restrictions to property apply in solving the conflict.[48] Therefore,

> 'States must take into account that indigenous territorial rights encompass a broader and different concept that relates to the collective right to survival as an organized people, with control over their habitat as a necessary condition for reproduction of their culture, for their own development and to carry out their life aspirations. Property of the land ensures that the members of the indigenous communities preserve their cultural heritage'.[49]

In fact, '[d]isregarding the ancestral right of the members of the indigenous communities to their territories could affect other basic rights, such as the right to cultural identity and to the very survival of the indigenous communities and their members.'[50] Conversely,

> 'restriction of the right of private individuals to private property might be necessary to attain the collective objective of preserving cultural identities in a democratic and pluralist society, in the sense given to this by the American Convention [on Human Rights]; and it could be proportional, if fair compensation is paid to those affected pursuant to Article 21(2) of the [American] Convention[on Human Rights].'[51]

This suggests that the different weight accorded to the interests at stake should normally lead to the primacy of indigenous peoples' right to property over the purely economic dimension of the 'individual' right to property of third parties. The protection afforded to the former is not absolute and states may not implement the restitution of the land 'for concrete and justified reasons', including the preservation of third parties' rights.[52] In these cases, 'the State must surrender alternative lands of equal extension and quality, which will be chosen by agreement with the members of the indigenous peoples, according to their own consultation and decision procedures.'[53] Once again, monetary compensation is not sufficient to settle land claims by indigenous peoples,[54] except

[45] Sawhoyamaxa Indigenous Community v. Paraguay, op. cit., para. 128. On the relevance of good faith, see section 5 below.
[46] Sawhoyamaxa Indigenous Community v. Paraguay, op. cit., para. 136.
[47] Ibid., para. 138.
[48] Ibid.; see also Yakye Axa Indigenous Community v. Paraguay, op. cit., paras 144-145.
[49] Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 146.
[50] Ibid., para. 147.
[51] Ibid., para. 148.
[52] Ibid., para. 149.
[53] Sawhoyamaxa Indigenous Community v. Paraguay, cit., para. 138. Cf. Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 149.
[54] It is noteworthy that this was the approach taken in the United States under the Indian Claims Commission Act of 1946; for a criticism, see Anaya, op. cit., p. 142.

DRAFT

'[w]here the peoples concerned express a preference for compensation in money or in kind, they shall be so compensated under appropriate guarantees.'[55]

5. ASSESSING THE CONDUCT OF THE OWNER/INVESTOR: THE RELEVANCE OF GOOD FAITH

As it was recently advocated by one scholar in the context of the balancing of private and public interests in the public utilities sector, the principle of proportionality should require the application of due diligence parameters to both the host state and the investor alike.[56] Indeed, in international investment law, the evaluation of the conduct of the investor seems to be confined to the assessment of domestic legality of the investment.[57] Interestingly, the case law, dealing mainly with corruption cases, affirms also a principle of estoppel, according to which:

> 'Principles of fairness should require a [investor-State arbitral] tribunal to hold a government estopped from raising violations of its own law as a jurisdictional defence when it knowingly overlooked them and endorsed an investment which was not in compliance with its law.'[58]

A common element emerging from the case-law of the two regional human rights courts is the relevance given, especially in determining entitlement to compensation, to the good faith of the person opposing her individual economic interest in property or investment to the common concern in the preservation of the environment.

Under the ECHR, the issue has been considered in cases where the authorities had ordered the destruction of buildings illegally erected in areas protected for their environmental[59] or cultural[60] value. At the outset, the question to be answered is whether such buildings qualify as 'possessions' in the autonomous meaning of Article 1 of Protocol No. 1. The ECtHR case-law demonstrates that good faith by the owner plays a pivotal role in deciding the matter.

In a case, the Court answered positively to this question noting that 'the applicants had acquired the property in question in good faith' and that 'they had peaceful enjoyment of their possession and had begun to build a hotel complex on the land as the lawful owners, after obtaining a building permit for that purpose'.[61] The fact that the construction was based on an administrative authorization attracted the protection of the

---

[55] Article 16, para. 4, of ILO Convention No. 169, quoted by the IACtHR in Yakye Axa Indigenous Community v. Paraguay, op. cit., para. 150.
[56] Tanzi, op. cit., p. 51.
[57] R. Dolzer and C. Schreuer, Principles of International Investment Law, New York: OUP, 2008, pp. 84-88; G. Bottini, 'Legality of Investments under ICSID Jurisprudence', in M. Waibel et al. (eds.) The Backlash against Investment Arbitration: Perceptions and Reality, Alphen aan den Rijn: Kluwer Law International, 2010, pp. 297-314; Viñuales, op. cit., pp. 158-261.
[58] Fraport AG Frankfurt Airport Services Worldwide v. The Republic of the Philippines, ICSID Case No. ARB/03/25, Award of 16 August 2007, para. 346.
[59] Hamer v. Belgium, op. cit., para. 76; N.A. and Others v. Turkey, no. 37451/97, Judgment of 11 October 2005 (Merits and Just Satisfaction).
[60] Vagnola S.P.A. and Madat S.R.L. v. Italy, no. 7653/04, Decision of 12 January 2010.
[61] N.A. and Others v. Turkey, no. 37451/97, Judgment of 11 October 2005 (Merits and Just Satisfaction), para. 39

DRAFT

ECHR. On the other hand, the Court has stated that the possession of a construction covertly built in non-compliance with planning regulations and which is set for demolition under domestic law does not per se constitute a protected 'possession' in the meaning of Article 1 of Protocol No. 1.[62]

However, the subsequent conduct of public authorities has to be considered in order to determine whether a proprietary interest in the enjoyment of the building concerned can be 'sufficiently established and weighty to amount to a substantive interest' protected.[63] The decisive element in this connection is whether the authorities have officially tolerated the illegal situation.[64] When the existence of a possession is ascertained, the respective conduct of the applicant and the respondent state are balanced in the proportionality assessment and the entitlement to compensation for the demolition of the illegally erected building seems to be limited to a situation where the behaviour of the authorities was such as positively engender in the applicant the conviction of being immune from successive law-enforcement actions.[65] Therefore, bad faith in gaining possession is a significant element in order to assess that a measure aimed at preserving the environment and the cultural heritage maintains proportionality, even in absence of compensation for the loss suffered by the concerned individual.

Good faith of third parties claiming 'private' property rights over traditional lands of indigenous communities is taken into account by the IACtHR in order to characterize the legal situation submitted to her. As discussed in the previous section, when the title to property or other rights over the lands have been lawfully transferred to 'innocent' third parties, the state is under the obligation to ensure their restitution. However, if those third parties did not acquire their titles in good faith, a concept upon which the IACtHR does not elaborate further, the property rights of the communities concerned are not extinguished and their entitlement to recover their traditional lands constitutes an aspect of the exercise of the right to property, rather than a restitution measure.[66] This difference plays an important role since in the latter case third parties' claims are not to be balanced against the restitution claims of indigenous peoples and no need for proportionality analysis arises. It is also implicit in this approach that the termination or revocation of titles acquired through open and illegal dispossession is not liable to compensation. Therefore, good faith of third parties owners plays a decisive role in distributing the financial burden of protecting indigenous people's territorial and environmental rights.

6. CONCLUDING REMARKS

[62] Vagnola S.P.A. and Madat S.R.L. v. Italy, op. cit.
[63] Hamer v. Belgium, op. cit., para. 76, and *Öneryıldız* v. Turkey [GC], no. 48939/99, Judgment of 30 November 2004 (Merits and Just Satisfaction), para. 129 and, mutatis mutandis, Depalle v. France, op. cit., para. 68.
[64] *Öneryıldız* v. Turkey, para. 127; Hamer v. Belgium, op. cit., para. 77.
[65] Vagnola S.P.A. and Madat S.R.L. v. Italy, op. cit.; Hamer v. Belgium, op. cit., para. 75.
[66] Sawhoyamaxa Indigenous Community v. Paraguay, op. cit., paras 127-128.

DRAFT

Far from suggesting an outright transposition of the human rights jurisprudence into international investment law and arbitration, the experience of regional human rights courts may at least indicate a reflection.

This is based on the assumption that, in legal terms, no real conflict exists between the right of states to regulate in order to attain progressively higher environmental standards, on the one hand, and the protection of the right to property (or of investments): in the worst scenario, under international investment law (and allegedly under customary law on expropriation) states are called to pay a 'prompt, adequate and effective' compensation for regulatory measures amounting to expropriation. In other words, the real issue at hand is how costs of increased environmental regulation should be allocated within a society.

A short analysis of human rights jurisprudence suggests that an approach based on a degree of deference to good faith sovereign decisions of states, on the balancing between competing interests that takes in account good faith of the owner/investor and on a flexible determination of compensation, in allowing different solutions tailor-made on the specific facts of each case, may tackle issues of 'distributive justice' much better than an 'all or nothing' approach.

Whether the reception of these concepts in the process of applying international investment law may take place through interpretation by arbitral tribunals (as the recent examples of the use of balancing may suggest) or whether new normative and/or institutional arrangements are need, is an issue for further discussion.

DRAFT

Check for updates

COLLOQUIUM INTRODUCTION

Role of economics in analyzing the environment and sustainable development

Stephen Polasky[a,b,1], Catherine L. Kling[c,d], Simon A. Levin[a,c,e], Stephen R. Carpenter[a,f], Gretchen C. Daily[a,g,h,i], Paul R. Ehrlich[a,g], Geoffrey M. Heal[j], and Jane Lubchenco[a,k]

The environmental sciences have documented large and worrisome changes in earth systems, from climate change and loss of biodiversity, to changes in hydrological and nutrient cycles and depletion of natural resources (1–12). These global environmental changes have potentially large negative consequences for future human well-being, and raise questions about whether global civilization is on a sustainable path or is "consuming too much" by depleting vital natural capital (13). The increased scale of economic activity and the consequent increasing impacts on a finite Earth arises from both major demographic changes—including population growth, shifts in age structure, urbanization, and spatial redistributions through migration (14–18)—and rising per capita income and shifts in consumption patterns, such as increases in meat consumption with rising income (19, 20).

At the same time, many people are consuming too little. In 2015, ~10% of the world's population (736 million) lived in extreme poverty with incomes of less than $1.90 per day (21). In 2017, 821 million people were malnourished, an increase in the number reported malnourished compared with 2016 (22). There is an urgent need for further economic development to lift people out of poverty. In addition, rising inequality resulting in increasing polarization of society is itself a threat to achieving sustainable development. Eliminating poverty (goal 1) and hunger (goal 2), achieving gender equality (goal 6), and reducing inequality (goal 10) feature prominently in the United Nation's Sustainable Development Goals (23). A recent special issue in PNAS on natural capital framed the challenge of sustainable development as one of developing "economic, social, and governance systems capable of ending poverty and achieving sustainable levels of population and consumption while securing the life-support systems underpinning current and future human well-being" (24).

The discipline of economics arguably should play a central role in meeting the sustainable development challenge. The core question at the heart of sustainable development is how to allocate the finite resources of the planet to meet "the needs of the present, without compromising the ability of future generations to meet their own needs" (25). A central focus of economics is how to allocate scarce resources to meet desired goals; indeed, a standard definition of economics is the study of allocation under scarcity. More specifically, economics studies the production, distribution, and consumption of goods and services, which are both a key driver of development (increasing standards of living through providing food, housing, and other basic human requirements) and a main cause of current changes in earth systems. Economics, combined with earth system sciences, is crucial for understanding both positive and negative impacts of alternatives and the trade-offs involved. Economics, combined with other social and behavioral sciences, is crucial for understanding how it might be possible to shift human behavior toward achieving sustainable development. Economics has well-developed fields in development economics, ecological economics, environmental economics, and natural resource economics, with large bodies of research relevant to the sustainable development challenge. The application of economic principles and empirical findings should be a central component in the quest to meet the aspirations of humanity for a good life given the finite resources of the earth.

Indeed, an extensive body of work by economists provides key insights into aspects of sustainable development. At its best, this work integrates work by

[a]The Beijer Institute of Ecological Economics, Royal Swedish Academy of Sciences, SE-104 05 Stockholm, Sweden; [b]Department of Applied Economics, University of Minnesota, St. Paul, MN 55108; [c]Resources for the Future, Washington, DC 20036; [d]Dyson School of Applied Economics and Management, Cornell University, Ithaca, NY 14853; [e]Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544; [f]Center for Limnology, University of Wisconsin–Madison, Madison, WI 53706; [g]Department of Biology, Stanford University, Stanford, CA 94305; [h]Woods Institute, Stanford University, Stanford, CA 94305; [i]Natural Capital Project, Stanford University, Stanford, CA 94305; [j]Graduate School of Business, Columbia University, New York, NY 10027; and [k]Department of Integrative Biology, Oregon State University, Corvallis, OR 97331

This paper results from the Arthur M. Sackler Colloquium of the National Academy of Sciences, "Economics, Environment, and Sustainable Development," held January 17–18, 2018, at the Arnold and Mabel Beckman Center of the National Academies of Sciences and Engineering in Irvine, CA. The complete program and video recordings of most presentations are available on the NAS website at www.nasonline.org/economics-environment-and.

Author contributions: S.P., C.L.K., S.A.L., S.R.C., G.C.D., P.R.E., G.M.H., and J.L. wrote the paper.

The authors declare no conflict of interest

Published under the PNAS license.

[1]To whom correspondence should be addressed. Email: polasky@umn.edu.

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.

other natural and social sciences into a policy-relevant framework and demonstrates the rich potential for collaborations among economists, natural scientists, and other social scientists on sustainable development challenges. For example, economists have developed integrated economic and climate models to address important climate change policy questions, such as how much and how fast greenhouse gas emissions should be reduced (26–31). In 2018, William Nordhaus shared the Nobel Prize in economics, in large part for his seminal work on such models. These models have sparked large debates within economics over fundamental issues such as the proper discount rate (32–35), and with the natural sciences over the likely scale of damages from climate change (36, 37). Another Nobel Prize winner in economics, Elinor Ostrom, used economic models to highlight the importance of governance and institutions for sustainable use of common property resources (38–40). Another important area of work by economists directly relevant to sustainable development defines and measures inclusive wealth (13, 41–49). Ken Arrow, yet another Nobel Prize winner in economics, was a leader in this field. It is also notable that the intellectual roots of inclusive wealth trace to work in the 1970s of two Nobel Prize winners in economics, William Nordhaus and James Tobin (50). Inclusive wealth is a measure of the aggregate wealth of society, including the value of natural capital along with the values of human capital, manufactured capital, and social capital. Inclusive wealth is a sufficient statistic for showing whether or not global society is on a sustainable trajectory. For the past two decades, the Beijer Institute of Ecological Economics, part of the Royal Swedish Academy of Sciences, has held annual meetings bringing together leading economists and ecologists to discuss issues at the intersection of ecology and economics, which have resulted in a number of high-impact papers (51). The idea for a forum to highlight work in economics on environment and sustainable development originated at one of these meetings.

Despite these examples and many others, the center of gravity in the analysis of sustainable development remains in the natural sciences, and the center of gravity in economics remains far removed from the challenge of sustainable development. The natural sciences that form the core of earth systems science, including ecology, geology, climatology, hydrology, and oceanography, are a logical place to start to build understanding of the current state and the evolution of earth systems. Natural scientists have taken the lead in prominent analyses of pathways to achieve sustainable development. For example, Pacala and Socolow (52) outline feasible methods using existing technology to reduce greenhouse gas emissions to secure a livable climate. Foley et al. (53) analyze how to meet growing food demand without expanding the footprint of agriculture. Costello et al. (54) suggest how extensive fishery reform could result in improved productivity and ecosystem health. Tallis et al. (55) analyze how to improve material standard of living for a growing population in ways that simultaneously sustain biodiversity, reduce greenhouse gas emissions, and reduce water use and air pollution. These works show that it is feasible to achieve multiple sustainable development goals with existing technology. The harder challenge is combining what is feasible in a biophysical sense with the difficult economic, political, and social hurdles that prevent society from getting to sustainable outcomes (55). In other words, natural science understanding is necessary but not sufficient to achieve sustainable development.

While natural science understanding is insufficient on its own to achieve sustainable development, the same is true of economics. Economists alone do not have the knowledge base supplied by the natural sciences necessary to understand the complex ecological systems within which the economic system operates and on which economic activity causes impacts. Progress in sustainable development requires collaboration between social scientists, including economists and natural scientists. Of course, achieving sustainable development requires institutions and political alignment that go well beyond assembling the science knowledge arising from integrated scientific knowledge.

Numerous examples show the incomplete nature of collaboration between economists and other disciplines engaged in the analysis of sustainable development. To take one recent example, there were no economists involved in a special section on "Ecosystem Earth" published in *Science* in April 2017 that contained discussions of population, consumption, agricultural production, land use, human behavior, collective action, and policy (56). The lack of involvement by economists in ongoing discussions of sustainable development leads to gaps in understanding production and consumption decisions, the resulting market outcomes that drive global environmental change, and how to regulate or reduce negative environmental impacts from economic activities.

The incomplete engagement of economists mirrors the structure of the economics discipline. The fields of ecological, environmental, and resource economics are not core fields within economics. There are few ecological, environmental, or resource economics publications in flagship journals within economics. For example, in 2018 only two papers published in the *American Economic Review* listed classification codes for renewable resources and conservation, nonrenewable resources and conservation, energy economics, or environmental economics (57, 58). Only a small minority of the top economics departments have fields in ecological, environmental, or resource economics. In contrast, virtually every top economics program offers fields in labor economics, industrial organization, and international trade. Ecological, environmental, and resource economics programs often are in schools of the environment or natural resources, schools of public policy, or in departments of agricultural economics. In addition, economics is notable among academic disciplines for its relative isolation: "Though all disciplines are in some way insular. . .this trait peculiarly characterizes economics" (59). Compared with other social scientists, economists have far lower citation rates for work in other disciplines. Jacobs (60) found that the percentage of within-field citations in economics was 81%, versus 59% for political science, 53% for anthropology, and 52% for sociology. In addition, the core of the economics discipline is relatively isolated from the natural sciences that have played a large role in sustainability science to date, ecology, geology, climatology, hydrology, and marine biology. Network maps of disciplines using citations patterns often show economics and fields, such as ecology and geosciences, at opposite ends of the spectrum (figure 3 in ref. 61).

Given the large role of economic activity in causing rapid change in earth systems, and the scale of the sustainable development challenge, there is an urgent need for more rapid integration of economics into the core of sustainable development, and for more rapid integration of sustainable development into the core of economics.

Sackler Colloquium on "Economics, Environment, and Sustainable Development"

This special issue contains a collection of articles presented at the Sackler Colloquium on "Economics, Environment, and Sustainable Development" held at the Beckman Center in Irvine, California in January 2018. The colloquium focused on 21st century challenges requiring advances in fundamental economics at the nexus of

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.

global environmental change and sustainable development. The main purpose of the colloquium was to highlight work by talented economists working in ecological, environmental, or resource economics on sustainable development challenges. Part of the motivation for the colloquium, and the publication of this collection of articles, is to convince economists that sustainable development challenges are vitally important to global society and pose intellectual challenges that are the equal of any subject currently analyzed by economists.

The symposium had four major themes: (*i*) ecosystem services and natural capital; (*ii*) behavioral economics, policy, and institutional design for sustainable development; (*iii*) economic development and sustainability; and (*iv*) issues in empirical economics relevant for sustainable development. Papers in this special issue address at least one, and in many cases several of these themes.

Ecosystem Services and Natural Capital. Ecosystem services are the contributions that nature makes to human well-being. Ecosystem services include regulating services (e.g., filtering pollution, coastal protection, pest regulation, pollination), material provisioning services (e.g., food, energy, materials), and nonmaterial services (e.g., aesthetics, experience, learning, physical and mental health, recreation). Various types of natural capital—often in conjunction with other forms of capital—and human labor provide ecosystem services. Destroying or degrading natural capital can result in reduced flow of ecosystem services, with consequent negative impacts on human well-being.

Research by economists, in conjunction with ecologists and other natural scientists, is essential for going beyond merely listing the types of ecosystem services and natural capital, to understanding the value of the flow of services or the stock in capital in terms of their contribution to human well-being. Integrated economic–ecological modeling can generate understanding of the trade-offs resulting from actions that alter ecosystems, and show how changes in ecosystems result in changes in the value of the flow of ecosystem services and the stock of natural capital. Examples of integrated work examining the value of ecosystem services and trade-offs exists at the national level (e.g., refs. 62 and 63) and local to regional levels (e.g., refs. 64 and 65), but much work remains to be done. Some of the pressing issues and questions that would benefit from greater involvement of economists are discussed below.

Measuring the value of ecosystem services. What are the best ways to apply market and nonmarket valuation methods to translate biophysical indicators into a common monetary metric measuring the welfare contribution of ecosystem services? One promising avenue to valuation links environment to health (66, 67). Difficult issues include integrating natural science and economic models to understand how changes in ecosystems lead to changes in the flows of ecosystem services (68, 69) and how to measure the value of nonmaterial ecosystem services, such as aesthetics, experience, learning, and mental health (70). Even where benefits measures exist, such as for improving water quality, it is not clear that all relevant benefits are accurately measured (71). Furthermore, it is often important to understand who benefits from ecosystem services and the distribution of benefits and costs of potential management and policy options.

Measuring the value of natural capital. Valuing natural capital involves making predictions about the future flows and values of ecosystem services (69). As baseball great Yogi Berra once said, "It's tough to make predictions, especially about the future." Recent efforts to estimate the social cost of carbon aptly illustrate the difficulties of valuing natural capital and ecosystem services. Estimating the social cost of carbon involves predicting future impacts, requiring integrated natural science and economic modeling, and understanding potentially catastrophic events (72). Constructing an estimate of the social cost of carbon also involves valuing nonmarket benefits, and classic economic problems, such as choosing an appropriate discount rate and degree of risk aversion (e.g., refs. 73 and 74). Economic analysis can increase knowledge of how various forms of capital, including natural capital along with human and manufactured capital, combine to produce goods and services of value to people (13, 44). To what degree can other types of capital substitute for natural capital, and what kinds of natural capital are irreplaceable?

Incorporating dynamics and uncertainty. Making predictions about earth systems is more difficult with the potential for tipping points involving large and sudden shifts (75). How can economic models incorporate notions of tipping points and the value of resilience and what data are needed to support empirical applications of such models (72, 76)? An additional issue arises with the combination of uncertainty and potential irreversible outcomes (e.g., species extinction) that gives rise to option value (77, 78). Issues are made more complex by the fact that most environmental–economic problems are characterized by ambiguity rather than risk, meaning that the familiar expected utility paradigm may not be applicable.

Behavioral Economics, Policy, and Institutional Design for Sustainable Development. Achieving sustainable development will require changes in human behavior and actions in relationship to the environment. Many important environmental and resources issues have elements of the "tragedy of the commons" where individuals following their own self-interest results in highly inefficient outcomes because individuals ignore external costs (or benefits) of their actions on others. Both standard and behavioral economics have much to offer in understanding what motivates individual and group behavior, how to structure incentives to shift behavior in desirable directions, and how to design policies and institutions to achieve desirable societal outcomes (79–81).

Some of the pressing issues and questions that would benefit from greater involvement of economists are discussed below.

Behavioral economics and individual choices involving environmental outcomes. What does behavioral economics teach us about trying to change behavior to overcome the tragedy of the commons, provide public goods, or internalize externalities? Are appeals to being good environmental stewards, information about performance relative to peers (e.g., energy or resource use relative to similar households), financial incentives, laws, and regulations more effective in promoting more proenvironmental behavior (82)? Do financial incentives crowd out nonfinancial motivations for protecting the environment or strengthen intrinsic motivation (82, 83)?

Social interactions and group behavior involving environmental outcomes. Humans are a social species. Economics has long studied individual behavior in isolation but there is ample evidence from social science that social interactions influence individual choices. How group interaction affects choices and environmental outcome, including cooperation to overcome the "tragedy of the commons," is a rich area of investigation (82, 84, 85).

Risk, uncertainty, and long-term consequences. How do people process risk, uncertainty, and ambiguity and what lessons does this hold for environmental issues that are inherently complex, with outcomes that are difficult to assign probabilities (76)? Are people myopic, and even if they are, can they be motivated to undertake current sacrifices to provide future generations with benefits?

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.

Design of environmental policy and institutions. Environmental issues span the gamut from quite local (e.g., communal use of a fishery, forest, or grazing lands) to global (e.g., climate change and ozone depletion) and involve complex feedbacks between social and ecological systems (86). Well-designed institutions can create incentives that drive performance toward desirable outcomes, or if ill designed can lead to poor outcomes (87). How can we design effective international environmental agreements for global environmental issues in a world of nation states? How can we effectively provide public goods and internalize externalities where governments are absent, weak, corrupt, or inefficient? When do payments for ecosystem services (88) or contributions to environmental groups (89) deliver desirable outcomes? What does the evidence from behavioral economics teach us about proper design of environmental policy and institutional design?

Economic Development and Sustainability. Sustainable development is not just about sustainability in the sense of how to maintain the environment. Sustainable development is about how to simultaneously alleviate poverty/improve material standards of living and maintain or enhance vital natural capital necessary for future well-being. Much of the work in environmental sciences focuses on environmental sustainability while much of development economics focuses on alleviating poverty. Making progress on sustainable development requires integration of research in development and environment.

Some of the pressing issues and questions that would benefit from greater involvement of economists are discussed below.

Integrating development and environment. What is the relationship between poverty alleviation and the environment? Are there win–win outcomes that reduce poverty and improve environmental quality (e.g., ref. 54) or are there inevitable trade-offs? Are payments for ecosystem services a promising approach for promoting environmental outcomes in developing countries (88)? Is the main effect of economic development to impose more stress on the environment through increased resource use and pollution from a greater scale of production and consumption, or are leap-frogging options available that will result in cobenefits vs. trade-offs? Will increased stress lead to crossing tipping points and catastrophic future outcomes, and how should society manage in the face of such prospects (72, 76)? Alternatively, does greater technological improvement and investment in pollution control, along with lower population growth that occurs at higher income levels outweigh the scale effect? What is the key role of protected areas in safeguarding biodiversity and potential adaptation to climate change vs. conversion of habitats to direct human uses? What different financing models might be developed to facilitate transitioning from unsustainable uses of natural resources to sustainable uses of those resources (e.g., fisheries), where the biggest financial impediment is making the transition, not sustaining the reformed fisheries?

Measures of sustainable development. Economists have developed measures of inclusive wealth that attempt to measure the value of all assets, including natural capital, human capital, and manufactured capital (13, 41–49). A difficult problem with empirical measures of inclusive wealth is correctly valuing natural capital for which there is no market value (see the above challenge on valuing ecosystem services and natural capital). At present, attempts of measure inclusive wealth include only a fraction of natural capital, typically only including natural capital values associated with natural resource commodities and carbon. Can we estimate comprehensive measures of inclusive wealth that include even difficult-to-value forms of natural capital?

Equity issues. Aggregate measures like inclusive wealth often hide distributional issues of who benefits and who does not from additional economic growth. Are there ways to represent distributional issues in a fair manner in aggregate economic measures like inclusive wealth or gross domestic product? Alternatively, is it better to disaggregate benefits and costs by group, and focus special attention on the poorest groups in society? If the world benefits from protecting habitat types in developing countries (e.g., tropical rain forests), what approaches would enable developing countries to protect these global public goods?

The role of innovation. Since the start of the Industrial Revolution, innovation and technology have fueled economic growth and rising material standards of living. However, because there is no price for most ecosystem services and natural capital, innovation incentives skew against maintaining or enhancing natural capital and the ecosystem services they provide. To promote proenvironment innovation, is it necessary to price ecosystem services and natural capital or, alternatively, is it better to provide direct incentives for innovation to enhance natural capital and ecosystem services?

Issues in Empirical Analysis. The past two decades have seen a major shift in empirical methods used by economists, with increased emphasis on research design using random assignment and quasi-experiments to better capture causal relationships (90). Although the "credibility revolution" was somewhat slower to take hold in environmental and resource economics than in some other fields, such as labor economics, recent years have seen an increased application of randomized control trials and quasi-experiments, particularly in applications involving individual or household decisions around use of land use, water, resources, and energy. In addition, the increasing availability of large, spatially detailed datasets along with increased computational power has opened the door to many new opportunities for rigorous investigation (91). Machine learning and other computationally rich approaches hold significant potential.

Some of the potential areas for innovation at would benefit from greater involvement of economists are discussed below.

Application of randomized control trials and quasi-experiments. Economists use randomized control trials and quasi-experiments to analyze individual or household decisions (92). A challenge is using such approaches to study larger-scale issues in sustainable development at regional, national, or international scales (86). There is a need for wider application of rigorous empirical methods for large-scale sustainable development issues.

Reduced form data intensive analyses. Use of large datasets with observations that vary across space and time offers an alternative approach to providing credible evidence on the impact of environmental policies when randomized control trials are either impossible or prohibitively expensive. The increase in geographically delineated data through time makes this approach applicable to a wider range of applications (66, 67, 93–95).

Integrated assessment models. The use of integrated assessment models has become well accepted in the context of greenhouse gas emissions and, increasingly, air-quality fate and transport (96), but needs exist for extensive application of these methods in other areas including water quality, water quantity, habitat, and so forth (97). Incorporation of both the drivers of change that can be impacted by policy so that costs of policy change can be included (63, 98), as well as downstream impacts, are also needed.

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.

Scenarios and future trends. Global assessments of environmental status and trends conducted by such entities as the Intergovernmental Panel on Climate Change, the Millennium Ecosystem Assessment, and the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services make use of scenarios to describe future trajectories. These scenarios are often based on expert opinion and do not have estimates about the likelihood that such a trajectory will occur. Can economists apply econometric tools and integrate economic data and principles with natural science to provide a more rigorous grounding on which to base future projections?

Summary

The challenge of achieving sustainable development is large and pressing. Major on-going changes in earth systems could cause potentially large negative consequences for human well-being. How to reduce poverty and address rising inequality in the face of these environmental changes raise important social and economic aspects of sustainable development in addition to the environmental dimensions. Strong arguments exist for devoting greater efforts to increase our understanding of the environmental, social, and economic dimensions of sustainable development, which will require greater integration of economics, social sciences more generally, and the natural sciences. Here, we focused on economics, but similar issues arise in other social and behavioral science disciplines. There is an urgent need for more rapid integration of social and behavioral disciplines into the core of sustainable development, and for more rapid integration of sustainable development into the core of these disciplines.

The papers in this special issue collectively represent a cross-section of work illustrating progress by economists on important issues in environment and sustainable development. These papers tackle many of the themes and challenges discussed in sustainable development and provide positive examples of the type of work needed to address sustainable development challenges. By providing a spotlight on the research frontiers, at least as they appear currently in economics, we hope to change the ecology of the economics profession in ways that stimulate further research and the involvement of economists in integrated research relevant to environmental and sustainable development challenges.

Acknowledgments

We acknowledge Partha Dasgupta for his help at the planning stage in setting major themes of the Sackler Colloquium. We dedicate this special issue to the memory of Kenneth Arrow and Elinor Ostrom, two intellectual giants who championed the integration of economics with other disciplines in pursuit of sustainable development.

1 Vitousek PM, et al. (1997) Human domination of earth's ecosystems. *Science* 277:494–499.
2 Millennium Ecosystem Assessment (2005) *Ecosystems and Human Well-Being: Synthesis* (Island Press, Washington, DC).
3 Ehrlich PR (2008) Key issues for attention from ecological economists. *Environ Dev Econ* 13:1–20.
4 Rockström J, et al. (2009) A safe operating space for humanity. *Nature* 461:472–475.
5 Intergovernmental Panel on Climate Change (2014) *Climate Change Synthesis Report: Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Intergovernmental Panel on Climate Change, Geneva).
6 Newbold T, et al. (2015) Global effects of land use on local terrestrial biodiversity. *Nature* 520:45–50.
7 Steffen W, et al. (2015) Sustainability. Planetary boundaries: Guiding human development on a changing planet. *Science* 347:1259855.
8 McRae L, Deinet S, Freeman R (2017) The diversity-weighted Living Planet Index: Controlling for taxonomic bias in a global biodiversity indicator. *PLoS One* 12:e0169156.
9 Díaz S, et al. (2018) Assessing nature's contributions to people. *Science* 359:270–272.
10 Intergovernmental Panel on Climate Change (2018) *Global Warming of 1.5°C: An IPCC Special Report on the Impacts of Global Warming of 1.5°C Above Pre-Industrial Levels and Related Global Greenhouse Gas Emission Pathways, in the Context of Strengthening the Global Response to the Threat of Climate Change, Sustainable Development, and Efforts to Eradicate Poverty* (World Meteorological Organization, Geneva).
11 Steffen W, et al. (2018) Trajectories of the earth system in the Anthropocene. *Proc Natl Acad Sci USA* 115:8252–8259.
12 Ripple WJ, et al. (2017) World scientists' warning to humanity: A second notice. *Bioscience* 67:1026–1028.
13 Arrow K, et al. (2004) Are we consuming too much? *J Econ Perspect* 18:147–172.
14 Ehrlich PR, Holdren JP (1971) Impact of population growth. *Science* 171:1212–1217.
15 Dasgupta PS, Ehrlich PR (2013) Pervasive externalities at the population, consumption, and environment nexus. *Science* 340:324–328.
16 Bongaarts J (2016) Development: Slow down population growth. *Nature* 530:409–412.
17 Bongaarts J, O'Neill BC (2018) Global warming policy: Is population left out in the cold? *Science* 361:650–652.
18 Seto KC, Güneralp B, Hutyra LR (2012) Global forecasts of urban expansion to 2030 and direct impacts on biodiversity and carbon pools. *Proc Natl Acad Sci USA* 109:16083–16088.
19 Godfray HCJ, et al. (2010) Food security: The challenge of feeding 9 billion people. *Science* 327:812–818.
20 Tilman D, Balzer C, Hill J, Befort BL (2011) Global food demand and the sustainable intensification of agriculture. *Proc Natl Acad Sci USA* 108:20260–20264.
21 World Bank (2018) *Poverty and Shared Prosperity 2018: Piecing Together the Poverty Puzzle* (World Bank, Washington, DC).
22 Food and Agricultural Organization, International Fund for Agricultural Development, UNICEF, World Food Programme, World Health Organization (2018) *The State of Food Security and Nutrition in the World: Building Climate Reslience for Food Security and Nutrition* (FAO, Rome).
23 United Nations (2015) *Transforming Our World: The 2030 Agenda for Sustainable Development. A/RES/70/1* (United Nations, New York).
24 Guerry AD, et al. (2015) Natural capital and ecosystem services informing decisions: From promise to practice. *Proc Natl Acad Sci USA* 112:7348–7355.
25 World Commission on Environment and Development (1987) *Our Common Future* (Oxford Univ Press, Oxford).
26 Nordhaus WD (1994) *Managing the Global Commons: The Economics of Climate Change* (MIT Press, Cambridge, MA).
27 Nordhaus WD, Boyer J (2000) *Warming the World: Economic Models of Global Warming* (MIT Press, Cambridge, MA).
28 Stern NH, et al. (2006) *The Stern Review: The Economics of Climate Change* (Cambridge Univ Press, Cambridge, UK).
29 Nordhaus WD (2008) *A Question of Balance: Weighing the Options on Global Warming Policies* (Yale Univ Press, New Haven, CT).
30 Stern N (2008) The economics of climate change. *Am Econ Rev* 98:1–37.
31 Nordhaus WD (2013) *The Climate Casino: Risk, Uncertainty, and Economics for a Warming World* (Yale Univ Press, New Haven, CT).
32 Nordhaus WD (2007) Critical assumptions in the Stern Review on Climate Change. *Science* 317:201–202.
33 Stern N, Taylor C (2007) Climate change: Risk, ethics, and the Stern Review. *Science* 317:203–204.
34 Weitzman ML (1998) Why the far-distant future should be discounted at its lowest possible rate. *J Env Econ Manage* 36:201–208.
35 Weitzman ML (2009) On modeling and interpreting the economics of climate change. *Rev Econ Stud* 91:1–19.
36 IPCC (2014) *Climate Change 2014: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (Cambridge Univ Press, Cambridge, UK).
37 Stern N (2016) Economics: Current climate models are grossly misleading. *Nature* 530:407–409.
38 Ostrom E (1990) *Governing the Commons: The Evolution of Institutions for Collective Action* (Cambridge Univ Press, Cambridge, UK).

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.

39 Ostrom E (2009) A general framework for analyzing sustainability of social-ecological systems. *Science* 325:419–422.
40 Ostrom E (2010) Beyond markets and states: Polycentric governance of complex economic systems. *Am Econ Rev* 100:641–672.
41 Dasgupta P (2001) *Human Well-Being and the Natural Environment* (Oxford Univ Press, Oxford).
42 World Bank (2006) *Where Is the Wealth of Nations?* (World Bank, Washington, DC).
43 World Bank (2011) *The Changing Wealth of Nations: Measuring Sustainable Development in the New Millennium* (World Bank, Washington, DC).
44 Arrow K, et al. (2012) Sustainability and the measurement of wealth. *Environ Dev Econ* 17:317–353.
45 United Nations University, International Human Dimensions Programme on Global Change, United Nations Environmental Programme (2012) *Inclusive Wealth Report 2012: Measuring Progress Toward Sustainability* (Cambridge Univ Press, Cambridge, UK).
46 Dasgupta P (2014) Measuring the wealth of nations. *Annu Rev Resour Econ* 6:17–31.
47 United Nations University, International Human Dimensions Programme on Global Change, United Nations Environmental Programme (2014) *Inclusive Wealth Report 2014: Measuring Progress Toward Sustainability* (Cambridge Univ Press, Cambridge, UK).
48 Polasky S, et al. (2015) Inclusive wealth as a metric of sustainable development. *Annu Rev Environ Resour* 40:445–466.
49 Managi S, Kumar P, eds (2018) *Inclusive Wealth Report 2018: Measuring Progress Towards Sustainability* (Routledge, Abingdon, UK).
50 Nordhaus WD, Tobin J (1972) Is growth obsolete? *Economic Research: Retrospect and Prospect Vol 5: Economic Growth*, eds Nordhaus WD, Tobin J (National Bureau of Economic Research, Cambridge, MA).
51 Söderqvist T, Sundbaum A, Folke C, Mäler K-G (2011) *Bringing Ecologists and Economists Together: The Askö Meetings and Papers* (Springer, Dordrecht, The Netherlands).
52 Pacala S, Socolow R (2004) Stabilization wedges: Solving the climate problem for the next 50 years with current technologies. *Science* 305:968–972.
53 Foley JA, et al. (2011) Solutions for a cultivated planet. *Nature* 478:337–342.
54 Costello C, et al. (2016) Global fishery prospects under contrasting management regimes. *Proc Natl Acad Sci USA* 113:5125–5129.
55 Tallis H, et al. (2018) An attainable global vision for conservation and human well-being. *Front Ecol Environ* 16:563–570.
56 Vignieri S, Fahrenkamp-Uppenbrink J (2017) Ecosystem Earth. *Science* 356:258–259.
57 Antolín-Díaz J, Rubio-Ramírez JF (2018) Narrative sign restrictions for SVARs. *Amer Econ Rev* 108:2802–2829.
58 Shapiro JS, Walker R (2018) Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade. *Amer Econ Rev* 108:3814–3854.
59 Fourcade M, Ollion E, Algan Y (2015) The superiority of economists. *J Econ Perspect* 29:89–114.
60 Jacobs JA (2013) *In Defense of Disciplines: Interdisciplinarity and Specialization in the Research University* (Univ of Chicago Press, Chicago).
61 Rosvall M, Bergstrom CT (2008) Maps of random walks on complex networks reveal community structure. *Proc Natl Acad Sci USA* 105:1118–1123.
62 Bateman IJ, et al. (2013) Bringing ecosystem services into economic decision-making: Land use in the United Kingdom. *Science* 341:45–50.
63 Lawler JJ, et al. (2014) Projected land-use change impacts on ecosystem services in the United States. *Proc Natl Acad Sci USA* 111:7492–7497.
64 Nelson E, et al. (2009) Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Front Ecol Environ* 7:4–11.
65 Goldstein JH, et al. (2012) Integrating ecosystem-service tradeoffs into land-use decisions. *Proc Natl Acad Sci USA* 109:7565–7570.
66 Tan Soo J-S, Pattanayak SK (2019) Seeking natural capital projects: Forest fires, haze, and early-life exposure in Indonesia. *Proc Natl Acad Sci USA* 116:5239–5245.
67 Williams AM, Phaneuf DJ, Barrett MA, Su JG (2019) Short-term impact of $PM_{2.5}$ on contemporaneous asthma medication use: Behavior and the value of pollution reductions. *Proc Natl Acad Sci USA* 116:5246–5253.
68 Polasky S, Segerson K (2009) Integrating ecology and economics in the study of ecosystem services: Some lessons learned. *Annu Rev Resour Econ* 1:409–434.
69 Adamowicz W, et al. (2019) Assessing ecological infrastructure investments. *Proc Natl Acad Sci USA* 116:5254–5261.
70 Daniel TC, et al. (2012) Contributions of cultural services to the ecosystem services agenda. *Proc Natl Acad Sci USA* 109:8812–8819.
71 Keiser DA, Kling CL, Shapiro JS (2019) The low but uncertain measured benefits of US water quality policy. *Proc Natl Acad Sci USA* 116:5262–5269.
72 Besley T, Dixit A (2019) Environmental catastrophes and mitigation policies in a multiregion world. *Proc Natl Acad Sci USA* 116:5270–5276.
73 Anthoff D, Tol RSJ, Yohe GW (2009) Risk aversion, time preference, and the social cost of carbon. *Environ Res Lett* 4:024002.
74 Greenstone M, Kopits E, Wolverton A (2013) Developing a social cost of carbon for US regulatory analysis: A methodology and interpretation. *Rev Environ Econ Policy* 7:23–46.
75 Lenton TM, et al. (2008) Tipping elements in the Earth's climate system. *Proc Natl Acad Sci USA* 105:1786–1793.
76 Anderies JM, Mathias J-D, Janssen MA (2019) Knowledge infrastructure and safe operating spaces in social–ecological systems. *Proc Natl Acad Sci USA* 116:5277–5284.
77 Arrow K, Fisher A (1974) Environmental preservation, uncertainty, and irreversibility. *Q J Econ* 88:312–319.
78 Henry C (1974) Investment decisions under uncertainty: The 'irreversibility effect.' *Am Econ Rev* 64:1006–1012.
79 Shogren J, Taylor LO (2008) On behavioral-environmental economics. *Rev Environ Econ Policy* 2:26–44.
80 Carlsson F, Johansson-Stenman O (2012) Behavioral economics and environmental policy. *Annu Rev Resour Econ* 4:75–99.
81 Kotchen MJ, Segerson K (2019) On the use of group performance and rights for environmental protection and resource management. *Proc Natl Acad Sci USA* 116:5285–5292.
82 Brandon A, List JA, Metcalfe RD, Price MK, Rundhammer F (2019) Testing for crowd out in social nudges: Evidence from a natural field experiment in the market for electricity. *Proc Natl Acad Sci USA* 116:5293–5298.
83 Frey BS, Jegens R (2002) Motivation crowding theory. *J Econ Surv* 15:589–611.
84 Hauge KE, Brekke KA, Nyborg K, Lind JT (2019) Sustaining cooperation through self-sorting: The good, the bad, and the conditional. *Proc Natl Acad Sci USA* 116:5299–5304.
85 Tilman AR, Dixit AK, Levin SA (2019) Localized prosocial preferences, public goods, and common-pool resources. *Proc Natl Acad Sci USA* 116:5305–5310.
86 Ferraro PJ, Sanchirico JN, Smith MD (2019) Causal inference in coupled human and natural systems. *Proc Natl Acad Sci USA* 116:5311–5318.
87 McDermott GR, Meng KC, McDonald GG, Costello CJ (2019) The blue paradox: Preemptive overfishing in marine reserves. *Proc Natl Acad Sci USA* 116:5319–5325.
88 Jack BK, Jayachandran S (2019) Self-selection into payments for ecosystem services programs. *Proc Natl Acad Sci USA* 116:5326–5333.
89 Grant L, Langpap C (2019) Private provision of public goods by environmental groups. *Proc Natl Acad Sci USA* 116:5334–5340.
90 Angrist JD, Pischke J-S (2010) The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *J Econ Perspect* 24:3–30.
91 Lemoine D, Rudik I (2017) Managing climate change under uncertainty: Recursive integrated assessment at an inflection point. *Annu Rev Resour Econ* 9:117–142.
92 Jessoe K, Lade GE, Loge F, Spang E (2017) Spillovers from behavioral interventions: Experimental evidence from water and energy use. E2e Working Paper 033. Available at https://e2e.haas.berkeley.edu/pdf/workingpapers/WP033.pdf.
93 Hollingsworth A, Rudik I (2019) External impacts of local energy policy: The case of renewable portfolio standards. *J Assoc Environ Resour Econ* 6:187–213.
94 Keiser D, Shapiro J (2019) Consequences of the Clean Water Act and the demand for water quality. *Q J Econ* 134:349–396.
95 Keiser D, Lade G, Rudik I (2018) Air pollution and visitation at U.S. national parks. *Sci Adv* 4:t1613.
96 Muller NZ, Mendelsohn R (2009) Efficient pollution regulation: Getting the prices right. *Am Econ Rev* 99:1714–1739.
97 Keiser D, Muller N (2017) Air and water: Integrated assessment models for multiple media. *Annu Rev Resour Econ* 9:165–184.
98 Rabotyagov SS, et al. (2014) Cost-effective targeting of conservation investments to reduce the northern Gulf of Mexico hypoxic zone. *Proc Natl Acad Sci USA* 111:18530–18535.

Downloaded from https://www.pnas.org by 37.101.222.112 on November 1, 2022 from IP address 37.101.222.112.