1. Introduction

International environmental problems have received increasing attention from economists in recent years. Two basic strands of literature may be distinguished. The first strand estimates the costs and benefits of various abatement targets under different cost allocation rules. It also discusses institutional issues and the design of treaties with respect to their efficiency. In particular the problem of global warming caused by the so-called greenhouse gases (see, for example, Brunner 1991; Cansier 1991; Chapman and Drennen 1990; Cline 1992a, b; Crosson 1989; Grubb 1989; IPCC 1996a, b; Manne and Richels 1991; Michaelis 1992; Nitze 1990; Nordhaus 1991a, b, c; Schelling 1991; and Welsch 1995) and the 'acid rain' problem due to sulfur and nitrogen oxides (for example, Crocker 1984; Førsund and Naevdal 1994; Foster 1993; Newbery 1990; Tahvonen *et al.* 1993; and Welsch 1990) have been studied in depth.

The second strand of literature has approached the problem of international pollution control from a game theoretical perspective (for example, Alho 1992; Andersson 1991; Buchholz and Konrad 1994; Chen 1997; Endres and Finus 1998a; Heister 1998; Kölle 1995; Kuhl 1987; Müller-Fürstenberger and Stephan 1997; Van der Ploeg and de Zeeuw 1992; and Welsch 1993). The incentive scheme of countries which sign a treaty and the stabilization of international environmental agreements (IEAs) are typical issues analyzed by this literature. This book is in the tradition of this second strand of literature.

Since, broadly speaking, game theory analyzes the interaction between agents and formulates hypotheses about their behavior and the final outcome in games, international environmental problems are particularly suited to analysis by these methods. Global and transboundary emissions exhibit a negative externality not only in the country of origin but in other countries too. Hence, there is a high interdependence between countries and strategic considerations enter the scene. Strategic aspects are particularly important in international pollution control since there is no 'world government' which could enforce IEAs. Therefore, the free-rider problem is a distinguishing feature of international environmental policy. Though countries are usually better off by coordinating their environmental policy, cooperation is often difficult to achieve. Since each country has only a marginal effect on aggregate emissions, it is always

better off letting others do the abatement job, thereby saving abatement costs.

Consequently, it is interesting to analyze the causes of this free-rider phenomenon and to derive conditions under which IEAs can, nevertheless, be stabilized. Thus, the central question of this book may be stated as follows: How can cooperation between countries to fight international pollution be established? The following analysis will focus on six aspects which crucially determine the prospects for cooperation. The first aspect concerns the *cost-benefit structure* of emission control. As a central result it will turn out that whenever cooperation generates high global welfare gains, IEAs achieve only little and are plagued by instability.

The second aspect is related to the *time dimension* of a game. Most generally, it will turn out that it is conducive to the stability of an IEA if IEAs are based on a long-term relationship between governments, if treaty obligations are regularly monitored and if instant reactions to a violation of a treaty are possible. In contrast, environmental projects suffer from high instability if they are connected with high sunk costs that involve a substantial time-lag to alter abatement strategies.

The third aspect concerns the *punishment options* in a game and the *credibility of threats*. Of course, the harsher the available punishments are, the easier it is to neutralize the free-rider incentive. However, it will turn out that, in the context of international environmental problems, punishments are normally not only detrimental to the punished but to the punishers as well. Hence, the question arises: Will a government be deterred by a threat of punishment which it believes will never be carried out since the punishers would hurt themselves? To use a metaphor: do we believe that a soldier would pull the pin out of a hand grenade if he will kill himself? The issue of the credibility of threat strategies will be given particular attention in this book. It is closely related to the definitions of equilibrium concepts. We shall discuss several equilibrium concepts which have emerged over recent years as refinements of the central concept in game theory: the Nash equilibrium.

The fourth aspect concerns the *enlargement of the strategy space* and is closely related to the term *issue linkage*. In reality it can frequently be observed that environmental issues are not negotiated in isolation but in connection with other policy fields. Concessions on one issue are traded against concessions on other issues. Such package deals may have several explanations. One is that they avoid possible asymmetric distributions of the gains from cooperation. Another is that issue linkage may ease the enforcement of an agreement. A government may threaten to withdraw from all agreements if one treaty is violated. This may increase the threat potential and thereby the stability of a treaty. It will be interesting to find

out in which situations issue linkage is conducive to cooperation (as is commonly believed) and whether there are conditions under which issue linkage should not be recommended to international negotiators.

The fifth aspect deals with the *institutional framework* in which negotiations take place. This issue evolves around the *instrumental choice* in international pollution control. We shall present a simple bargaining game in which governments bargain on the level of a uniform emission reduction quota and a uniform effluent charge. It will be shown in various chapters that the commonly believed superiority of a tax (market-based instrument) over a quota regime (command and control instrument) no longer holds in a second-best world where this second-best world is constituted by the following restrictions: the accession to an IEA must be voluntary; governments settle for the lowest common denominator proposal and agree on that institutional framework which a majority of governments favor; and the stability of an IEA must be enforced. The model provides many reasons for the fact that emission reduction quotas are part of most of the IEAs signed so far, but as yet no effluent charge has been applied in international pollution control, although economists strongly advocate market-based instruments on efficiency grounds.

The sixth aspect is concerned with *coalition formation* in international pollution control. The central three questions to be answered are: (a) how many and which countries sign an IEA? (b) will there be one or several coexisting agreements? and (c) how effective will agreements be?

The answers turn out to be highly complex since the set of possible government strategies in an N -country world is almost infinitely large. This is particularly true in negative externality games since the strategies of a coalition depend on the strategies of all other coalitions, which in turn depends on the overall coalition structure in the game. The final objective is to explain the entire coalition formation process endogenously.

The subsequent analysis assumes transboundary or global emissions which are summarized under the term *international environmental problems*. Typical examples of *transboundary pollutants* include the aforementioned acid rain or the salination of the river Rhine due to the potash mines in France (upstream country) from which Germany (downstream country) suffers. The first example is a *multilateral externality*, the second example a *unilateral externality*. Typical examples of *global pollutants* include the aforementioned greenhouse gases or substances which deplete the ozone layer.¹

Transboundary pollutants are also classified as *impure public bads* since the distribution of pollutants is usually uneven across countries. Accordingly, global pollutants belong to the group of *pure public bads* since emissions mix uniformly in the atmosphere and all countries suffer from the

externality approximately equally.² According to this classification it comes as no surprise that, apart from the environmental economics literature, the literature on the economics of public goods has also investigated the problem of cooperation and free-riding in the provision of public goods, though this literature does not explicitly refer to international environmental problems. The provision of a public good, say x , depends on the sum of contributions of each agent i , that is, $x = \sum x_i$, and agent i derives utility from x , that is, $u_i(x)$. If x_i is interpreted as abatement or emission reduction from some status quo, then it is obvious that a public goods model can be used to describe the structure of international environmental problems studied in this book. Therefore, some of the approaches in the public goods literature to explain the voluntary provision of public goods will be investigated in Chapter 10 in a game theoretical context.

The basic situation which will be analyzed subsequently may be described as follows. Two or more countries emit a transboundary or global pollutant. They are currently not cooperating but are considering doing so. Each government pursues its interests non-cooperatively, that is, it behaves individually rationally. This implies that governments only cooperate if it is in their interest, and that they always take a free-ride whenever this seems profitable to them. Implicitly, we assume that governments maximize some kind of welfare function that measures the gains from emissions accruing from the production and consumption of goods and the losses of emissions accruing from environmental damages. Thereby, the aggregation of welfare is treated as a black box. That is, we abstract from problems mentioned in the public choice literature concerning the interaction of interest groups, voters and bureaucracy and their effect on governmental decisions (Endres and Finus 1996a, b; Ursprung 1992).

Since there is no international agency which could enforce an agreement among sovereign countries, any IEA must be *self-enforcing*. The methods used to analyze such a situation will be taken from game theory, particularly from non-cooperative game theory.³ The approach followed in this book thereby strongly adheres to the principles of neoclassical economics.

This kind of reasoning has often been attacked as being unrealistic. In particular, the postulate of rationality has been subjected to severe criticism. The critics question whether human beings are as rational as game theory typically assumes them to be. However, this critique has also been raised against neoclassical economics in general. Since arguments in defense of the postulate of rationality are laid down in mighty words elsewhere, it is not necessary to restate them again here. We would only like to point out that we believe that governments approximately act in a rational manner. One can, however, disagree about the appropriate arguments which should be part of governments' objective function.

Another popular charge against game theory may be summarized in the following slogan: ‘With game theory one can prove *any* hypothesis – and its antithesis at the same time, too.’ Although this slogan has the charm of being humorous, it nevertheless is false. In game theory, as in any other theory, assumptions drive results. Sticking to assumptions once made, one will never – not even in game theory – find contradictory results. If the slogan is interpreted in a friendlier way, it can be seen as a claim that assumptions should be chosen very carefully; that is, they should be based on some plausibility and should, ideally, be empirically supported. Obviously, this (pretty trivial) claim is not only valid for game theory but applies to all theoretical reasoning.

To refute this last general critique, this book will strongly emphasize the underlying assumptions of the models analyzed. Moreover, the way in which they drive the results will be revealed explicitly. With a ‘critical distance’, the assumptions will be evaluated with respect to their aptness to reflect a particular problem. An overview of the most important assumptions is given in Chapter 2, Section 2.3.

There are two types of game theoretical literature with respect to the mathematical level of the presentation. The first type is highly technical, which makes it difficult for the mathematically less interested reader to access the material. This literature has dominated the game theoretical scene for a long time, which probably explains the fact that widespread applications to economics have only occurred within the last twenty years, though the roots of game theory can be traced back to J. von Neumann and O. Morgenstern (1944). The second type of literature could be responsible for the rapid expansion of applications of game theory to economics in recent years.⁴ It explains the central results of game theory intuitively but applies less rigor to formal proofs. We draw attention in particular to the books of Eichberger (1993); Gibbons (1992); Holler and Illing (1993); Kreps (1990); Myerson (1991); and Rasmusen (1995). Simple applications to environmental economics may be found in Burger (1994); Fees (1995); and Weimann (1995).

This book is intended to bridge the gap between these two strands of literature. It starts from the basics of non-cooperative game theory (Chapter 3) and then progresses step by step to the most advanced topics of coalition formation of recent years (Chapter 15). It is aimed at readers with a basic knowledge of microeconomics and mathematics⁵ and virtually no knowledge of game theory. Therefore, Chapter 2 introduces the reader to the most important terms used subsequently; other terms will be introduced in the course of reading the book. The proofs of most propositions are provided either in the text or in appendices: exceptions are proofs which are either obvious or would need too much space. Though the intuition of

all proofs is emphasized, those readers who are less interested in the technical part of a proof should be able to understand the central idea of all results from reading the text. Thus, this book has the, surely, heroic aim of being self-contained.

In what follows, Chapter 2 introduces some important terms of game theory (Section 2.1) and some frequently used notations in this book (Section 2.2). The second chapter also provides a taxonomy of game theory (Section 2.3). In the course of the discussion of this taxonomy it will be emphasized which parts of game theory are covered in this book. Therefore, the outline of the book has been left to Section 2.4.

NOTES

1. For a complete classification of environmental problems, see Siebert (1985, 1992).
2. However, this does not imply that environmental damages are equally perceived and evaluated.
3. The details of this setting and the terms mentioned will be explained in Chapter 2.
4. In the environmental economics context, see, for instance, the volumes edited by Hanley and Folmer (1998) and Pethig (1992).
5. Though some proofs may be very long and therefore may look tricky, basic algebra is sufficient to follow all the proofs.