Playing a Better Global Warming Game: Does it Help to be Green?

ALFRED ENDRES* and MICHAEL FINUS*

1. INTRODUCTION

It is a distinguishing feature of many international environmental problems that agreements on pollution reduction are difficult to obtain. Moreover, they are potentially unstable. Even if an agreement is designed to benefit all signatories, there usually prevails a strong incentive for the participants to take a free-ride on other countries' abatement efforts (BARRETT 1994 and CARRARO/SINISCALCO 1992). This prisoners' dilemma type of situation is particularly unfavorable: Different from national environmental problems where a government can enforce environmental policy, no similar agency exists at the global level.[1]

Therefore, it is important to identify the factors potentially improving the chances for agreements to be signed and enforced. In the policy debate two factors are frequently mentioned: environmental awareness and unilateral actions. It is commonly believed that an increase in environmental awareness is conducive to the protection of the global commons. Moreover, it has been proposed that governments should «give a good example» by unilaterally reducing emissions (HOEL 1991).[2] Unilateral action might happen before or after the agreement: On the one hand, if a country reduces emissions before a treaty is negotiated, this is a call upon its neighbors to follow suit and to approve more ambitious abatement targets. On the other hand, if a country reduces emissions after a convention has been signed, this may compensate for low abatement targets, which are often due to a settlement on the lowest common denominator in international politics. Indeed, the logic behind these arguments seems compelling and the intuition is straightforward.

* University of Hagen, Germany.
The authors gratefully acknowledge the receipt of grant No. II 69 982 by the Volkswagen-Foundation, Germany. The paper benefited from discussions at the Symposium on «Responsible Environmental Behavior», held at the IKAÖ, University of Bern, September 4.–7., 1996. The authors are indebted to URS LUTERBACHER, GUNTER STEPHAN and several other participants for helpful comments and suggestions. The authors also like to thank an anonymous referee for constructive comments on an earlier draft.

1. The International Court of Justice can only deal with non-compliance if the party accused of some wrong doing accepts to open a trial. Thus, formally, this court does not have any enforcement power (BOTHE 1995).
2. Possibly, this issues are interrelated. Increasing environmental awareness in a country might put some pressure on a government to act unilaterally.

However, good intentions are sometimes not enough, particularly if political representatives behave strategically. To demonstrate this, we use a simple analytic framework, abstracting from transaction costs and assuming a game of complete information.[3]

The paper proceeds as follows: In chapters 2 and 3 the effects of environmental consciousness on the stability of an agreement are analyzed in a static and a dynamic context.[4] In chapter 2 we lay out the basic model and investigate the effects for a «socially optimal» agreement. Chapter 3 extends the analysis to a more pragmatic solution, i.e., a uniform emission reduction quota. In chapter 4 the effects of unilateral abatement efforts by a country are discussed in a bargaining framework on emission quotas. Chapter 5 summarizes the main findings and draws some conclusions.

2. THE ROLE OF ENVIRONMENTAL AWARENESS IN A SOCIALLY OPTIMAL AGREEMENT

2.1. *Welfare Implications*

Assume n countries suffering from a global pollutant. Hence, damages in a country i depend on aggregate emissions. Moreover, let damages in a country, D_i, increase in aggregate emissions ($D_i' > 0$) at an increasing rate ($D_i'' > 0$). This is the standard assumption in the literature and reflects the fact that self-purification of environmental systems becomes increasingly difficult at high levels of pollution. Moreover, if environmental costs are monetized, these costs may be measured as the willingness to pay for emission reductions.[5] Then, it seems plausible (and empirically sound) to assume that the marginal willingness to pay is higher at higher levels of pollution than at lower levels.[6]

However, emissions do not only cause damages to the environment, they also generate utility to society via the consumption and production of goods. Thus, emissions E_i may be viewed as an input in the production and consumption process. Again, following a standard type of argument, we assume benefits derived from emissions B_i to increase in emission levels in some range ($B_i' \geq 0 \, \forall \, E_i \in \left[0, E_i^U\right]$), but at a decreasing rate ($B_i'' < 0$).[7]

3. The analysis is confined to the traditional neoclassical approach. For an alternative analytical framework see MOHR (1994). For a discussion of the implications if these assumptions do not hold, see MÜLLER-FÜRSTENBERGER/ STEPHAN (1996).
4. See also ENDRES (1997b) on the role of environmental awareness. There, the analysis is confined to two countries and a static context.
5. For the concept, methods and problems of monetization see, e.g., ENDRES/HOLM-MÜLLER (1998), FINUS (1992) and FREEMAN (1993).
6. See e.g. DESVOUGES et al. (1989) and HANLEY/SPASH (1993), pp. 216–218.
7. The assumption of decreasing marginal benefits of pollution is equivalent to the assumption of increasing marginal costs of pollution abatement. The later interpretation is also standard in the literature. See any environmental economics textbook, e.g., HANLEY et al. (1997) and JAEGER (1994).

To depict these relations, we choose the following simple net benefit function ($NB_i = B_i - D_i$)[8]

$$NB_i = b(dE_i - \frac{1}{2}E_i^2) - \frac{\alpha_i}{2}(\sum_{j=1}^{n} E_j)^2 \text{ with } i, j \in \mathbb{N} \text{ and } b > 0; d > 0; \alpha_i > 0. \quad (1)$$

Here, b, d and α_i are positive parameters. The subscript i refers to the country. Obviously, we have to require $E_i^U \leq d$ for the upper bound of the emission space to ensure $B_i' \geq 0$. Below, it will turn out that this condition is not binding since we define the strategy space of the global emission game to be $E_i \in [0, E_i^N]$, where E_i^N are emissions in country i in the Nash equilibrium of the stage game.

The parameter α_i is taken as a proxy variable to represent environmental awareness in a country i. This parameter values the damages from emissions. The higher the value of α_i, the more weight is given to emission reductions.[9] Typically, the value of this parameter will differ between countries. This implies that countries hold different views regarding the extent of emission control which makes it difficult to agree on joint abatement measures. The Rio Conference and its follow-ups as well as the ongoing discussion on joint implementation to control greenhouse gases are prominent examples of this phenomenon.

Due to these difficulties, joint abatement policies often fail and countries choose their non-cooperative emission levels. If all countries follow such a strategy, equilibrium emissions are given by the Nash equilibrium of the stage game. For (1) this equilibrium is unique and emissions are given by

$$E_i^N = \frac{d(b+\alpha-n\alpha_i)}{b+\alpha}; \; E^N = \frac{nbd}{b+\alpha} \text{ with } \alpha = \sum_{i=1}^{n} \alpha_i . \quad (2)$$

Since negative emissions have no economic meaning, we have to require $C_1 := b + \alpha \geq n\alpha_i$. This condition will be useful to establish some of the results below. Casual inspection of (2) reveals that $E_i^N < d$ holds, as claimed above.

A «socially optimal» solution would require that each country does not only take its own damages into consideration when choosing its emission level, but also those in the $n - 1$ other countries. Hence, instead of Max. $(B_i - D_i)$, a country performs Max. $(B_i - \sum D_i)$ in the social optimum which leads to emissions

8. Similar functions have been used by BARRETT (1994) and HOEL (1992).
9. We implicitly assume that a government (its representatives) maximizes welfare in the traditional way (the sum of its citizens' welfare). Hence, we discard all problems associated with the transformation of voters' opinion into a goverment's objective function (see BERNHOLZ/BREYER 1994 and ENDRES/FINUS 1996). Taking a public choice perspective would certainly be interesting but is beyond the scope of this paper.

$$E_i^S = \frac{bd}{b+n\alpha}\,;\; E^S = \frac{nbd}{b+n\alpha}\,. \tag{3}$$

Here, the superscript «S» stands for «social optimum». Comparing these emission levels with the emissions in the Nash equilibrium we derive a standard result: Emissions in the Nash equilibrium are too high from a global perspective, i.e., $E_i^N > E_i^S \; \forall i$, and therefore of course $E^N > E^S$.

Inserting the emission levels of (2) and (3) into (1), we derive net benefits. The following results can be established:

$$\left.\frac{\partial\left(NB_i^S - NB_i^N\right)}{\partial\alpha_i}\right|_{\alpha=const.} \geq 0,\; NB_i^S \lesseqgtr NB_i^N\,;\; NB^S > NB^N\,; \tag{4}$$

That is, countries with a higher environmental awareness benefit more from a socially optimal agreement than those with a lower consciousness. In fact, a country with a low environmental awareness does not necessarily gain in the social optimum compared to the Nash equilibrium, though aggregate welfare is higher. Whenever environmental awareness differs between countries, those countries with low values of α_i may be worse off in the social optimum than in the Nash equilibrium. The exact relations depend on the parameter values of the model.

The reason for this result is simple: Compared to their «low interests» in global pollution control, countries with low values of α_i have to contribute relatively much to the joint abatement policy in the social optimum. If we assume for instance a discrete uniform distribution of the parameter α_i in the interval $[1, n]$ (see, e.g., FINUS/RUNDSHAGEN 1997 and HOEL 1992), then it turns out that for the country with lowest index (country 1) $NB_1^S < NB_1^N$ holds, irrespectively of the parameter values. More general with respect to (1), one can show that a sufficient condition for $NB_i^S < NB_i^N$ is $\alpha_i < \alpha / n^2$. Thus, whenever the variance of the α_i-distribution is high, some countries will find it not attractive to join a socially optimal agreement. This contributes to understand why international environmental agreements usually specify more moderate emission reductions and distribute abatement burdens more equally, if they are ratified at all. This issue will be taken up in Section 3 in more detail.

2.2. Stability in a Static Context

In this section we assume that a socially optimal treaty has been signed and that countries comply with their abatement obligations, in the first place. In order to find out how stability of such an agreement is affected by a change in environmental awareness, we define an «incentive to defect» $I_i = NB_i^D(E_i(E_{-i}^S), E_{-i}^S) - NB_i^S(E_i^S, E_{-i}^S)$, where «$-i$»

denotes all other countries except i, and $E_i\,(E^S_{-i})$ is the «best reply» of country i to the emission levels of the other countries, through which country i nets a free-ride payoff NB^D_i. Thus, $NB^D_i(E_i\,(E^S_{-i}), E^S_{-i})$ is country i's maximal attainable net benefit from defecting while other countries comply with the treaty. The following interesting relations can be established:

$$I_i > 0; \frac{\partial I_i}{\partial \alpha_i} < 0; \frac{\partial I_j}{\partial \alpha_i} > 0; i \neq j. \tag{5}$$

In a socially optimal agreement there exists an incentive to free-ride for each country. This is not surprising since none of the elements of the socially optimal emission vector is located on a country's reaction function. From this it follows that a socially optimal agreement cannot be stable in a static context. Therefore, below, we will check for stability in a dynamic context. For the moment, however, we discard this problem.

If environmental awareness increases in a country i, this will lower its own free-rider incentive. However, it will increase those of its neighbors. To get the intuition behind the latter result consider that an increase in environmental awareness implies that more ambitious abatement targets are specified in a treaty. [This is obvious when observing (3)]. So after an increase in i's environmental awareness, the socially optimal emission vector is more distant to each country's reaction function (other than i's) than before. Hence, the «incentive to move back to their reaction functions» becomes stronger for these countries. This explains the detrimental effects of a rise in environmental consciousness in country i on the contractual fidelity of all other countries.

With respect to the effect of i's environmental awareness on i's own incentive to defect the same forces are at work that have been explained above with respect to the countries other than i. However, for country i, a second factor has to be taken into consideration. An increase in α_i shifts country i's reaction function more towards the origin and reduces the slope of this function (in absolute terms) in the n-dimensional emission space. Hence, due to an increase in α_i, country i would choose a lower emission level in the Nash equilibrium and its reaction function is «more sensitive» to emissions in the other countries. This effect works in the direction of a lower incentive to free-ride. Obviously, this second effect dominates the first effect explained above.

An other interesting result is obtained if we assume that environmental awareness is raised simultaneously in all countries. To study this effect we define $\alpha = \alpha'\gamma$ and $\alpha_i = \alpha_i'\gamma$, where γ can be interpreted as a parameter representing «general environmental awareness» in all countries. Substituting the equations above into the index to defect I_i, and differentiate I_i with respect to γ. We get (after $\alpha = \alpha'\gamma$ and $\alpha_i = \alpha_i'\gamma$ have been substituted back)

$$\frac{\partial I_i}{\partial \gamma}=\frac{b^2d^2n^2(\alpha_i-\alpha)^2(2b^2+b\alpha_i-\alpha\alpha_i n)}{2(b+\alpha_i)^2(b+\alpha_i n^2)^3}<0 \text{ iff } \alpha>\frac{b(2b+\alpha_i)}{\alpha_i n}. \tag{6}$$

Thus, only if «total environmental awareness» has been raised sufficiently high, the incentive to defect decreases with an increase in α. For low levels of α, a simultaneous (and proportional) increase in environmental consciousness in all countries leads to an even higher free-rider incentive. This result is striking and can be explained by noting the following relation:

$$\frac{\partial(E^N-E^S)}{\partial\alpha}=\frac{nbd(n-1)(b^2-n\alpha^2)}{(b+\alpha)^2(b+n\alpha)^2}<0 \;\; \textit{iff} \;\; \alpha>\frac{b}{\sqrt{n}} \tag{7}$$

The difference between emissions in the Nash equilibrium and the social optimum increases in α for low levels of environmental awareness. If a certain level of α is reached, increasing environmental awareness further will narrow the gap between both emission levels. In this second parameter range, socially optimal emission levels are relatively low so that further emission reductions – due to a raise in environmental awareness – cause marginal opportunity costs of abatement to increase substantially. Finally, opportunity costs become a binding factor for further emission reductions. This relation is an immediate consequence of the assumption of a concave net benefit function as specified in (1) and applies for cooperative and non-cooperative emission levels. Therefore, the limit of $(E^N - E^S)$ approaches zero for α going to infinity.

Since, as a tendency, the more demanding abatement targets are specified in a treaty with respect to the Nash equilibrium, the higher are the free-ride incentives of countries, the relation in (7) explains the forces at work generating the result in (6).

2.3. Stability in a Dynamic Context

In this section we test for stability in a supergame framework. We restrict attention to subgame-perfect strategies, though more sophisticated strategies have been applied in the literature (see, e.g., BARRETT 1994, ENDRES/FINUS 1998 and FINUS/RUNDSHAGEN 1996). This allows to investigate the stability issue analytically without having to rely on simulations.

For this class of strategies it is true that if the strategy space is restricted to $E_i \in [0, E_i^N]$, i.e., the harshest punishment for a free-rider is if all players play their Nash equilibrium stage game strategies for the rest of the game. This determines the lowest discount factor requirements in a stable agreement – i.e., the rates countries can use to discount their payoff stream over the entire game and compliance can still be enforced by subgame-perfect threats.

From the definition of the strategy space it follows that only agreements are stable which generate a payoff to each country at least as high as in the Nash equilibrium of the stage game. Hence, a socially optimal agreement might not be stable even for a discount factor close to one if countries are different with respect to environmental awareness (see Section 2.1). Therefore, we restrict attention to symmetric countries ($\alpha_i = \alpha_j$) in this section.

A socially optimal agreement is said to be a subgame-perfect equilibrium if it satisfies

$$\frac{NB_i^S}{1-\delta} \geq NB_i^D + \frac{NB_i^N \delta}{1-\delta} \text{ or } \delta \geq \frac{NB_i^D - NB_i^S}{NB_i^D - NB_i^N} := \delta_i^{Min}, \tag{8}$$

where δ_i^{Min} denotes the «minimum discount factor» of country i. Thus for $\delta_i^{Min} \leq \delta_i < 1 \, \forall \, i$ the agreement can be controlled using the harshest trigger strategy which is permissible in the global emission game. Differentiating δ_i^{Min} with respect to environmental awareness reveals the following relations:

$$\left.\frac{\partial \delta_i^{Min}}{\partial \alpha_i}\right|_{\alpha_i=\alpha_j} < 0; \left.\frac{\partial \delta_j^{Min}}{\partial \alpha_i}\right|_{\alpha_i=\alpha_j} > 0;$$

$$\frac{\partial \delta_i^{Min}}{\partial \alpha} = \frac{b(n-1)^2(b+n\alpha_i)(n\alpha_i - b)}{\left(\alpha_i n^2 b + 2\alpha_i^2 n^2 + 2bn\alpha_i + b\alpha_i + 2b^2\right)^2} < 0 \quad iff \quad \alpha_i < \frac{b}{n} \, \forall \, i, \text{ or } \alpha < b \tag{9}$$

Again, raising environmental awareness in country i reduces the stability requirements in this country, but has a detrimental effect on the other countries.[10] This result corresponds to the findings in the static context. Since the nominator of δ_i^{Min} is equal to I_i in (5), this result can be basically put down to the same effects as described in Section 2.2.

What is different now is the effect of raising environmental awareness simultaneously in all countries on stability. For low levels of α an increase in environmental awareness reduces the stability requirements (δ_i^{Min} decreases), whereas for high values this relation is reversed. The intuition is the following: As noted above, for higher values of α a country conducts a considerable amount of abatement even in the absence of an international agreement. The difference between the non-cooperative and the socially optimum emission level shrinks for higher values of environmental awareness [see (7)]. This implies that the threat to punish non-compliance by returning to the Nash equilib-

10. More optimistic results regarding the stability of an agreement in a dynamic setting could be obtained by assuming that the actual discount factor of countries is endogenously determined by environmental awareness. (We like to thank GUNTER STEPHAN for raising this point). For an endogenous treatment of the discount factor in the environmental context see HOWARTH/NORGAARD (1992) and STEPHAN/ MÜLLER-FÜRSTENBERGER/PREVIDOLI (1997).

rium of the stage game becomes less deterrent for a potential free-rider. So at high values of α, increasing environmental awareness reduces the severeness of the punishment.

Obviously, the «deterrence effect» dominates the «incentive effect» described in Section 2.2.

Taken together the results derived in Section 2 indicates that sometimes «good intentions» may have «counter-intuitive» (and unintended) effects. An increase in environmental awareness in a country or at the global level might increase the stability requirements for some participants in an agreement. If these forces are strong enough, it might be impossible to control an agreement. Of course, one might argue that socially optimal agreements will hardly be observed in reality and therefore these findings might not carry over to the analysis of practical problems. To check this, stability of a «typical» real world agreement, i.e., a convention on a uniform emission reduction quota is investigated in the next section.

3. THE ROLE OF ENVIRONMENTAL AWARENESS IN AN AGREEMENT ON A UNIFORM EMISSION REDUCTION QUOTA

3.1. Preliminaries

Looking at the history of international agreements one finds frequently that abatement targets are specified with respect to uniform emission reduction quotas. This is to say that all signatories must reduce their emissions by the same percentage to some base year. For instance the Montreal Protocol specifies a 20 percent emission reduction of CFCs in all countries based on 1989 emission levels to be accomplished by 1998. An other example is the Helsinki Protocol which suggests a reduction of long range transboundary air pollutants from 1980 levels by 30 percent by 1993.

The frequent appearance of uniform solutions seems to be due to the notion in public that these solutions are «fair» or «just». Moreover, to agree on differentiated solutions takes time and is therefore associated with higher transaction costs. Hence, uniform solutions constitute some kind of a focal point in the sense of SCHELLING (1960) to which bargaining partners feel relatively easy to oblige (BARRETT 1992 and HOEL 1992).

With respect to the particular prominence of emission reduction quotas compared to other policy instruments, e.g. emission tax or tradable permits, there are several explanations presented in the literature. It has been argued that equal percentage emission reductions require the participating countries to possess less information compared to negotiations on effluent charge rates (ENDRES 1996, 1997a). In particular, neither country needs to know other countries' marginal damage costs. All what is required is to assume that each country has information on current emission levels in all countries (and its own damage and abatement costs). Moreover, it has been shown that in a second best world agreements on quotas may even be Pareto-superior to taxes and tradable permits if countries exhibit certain asymmetries (ENDRES 1996, 1997a and ENDRES/FINUS 1998).

It has also been demonstrated that the discount factor requirements for the countries with a low interest in a joint abatement policy are usually lower in a quota than in an emission tax regime. This is particularly true if countries have very different interests (FINUS/RUNDSHAGEN 1996). Finally, in a model on coalition formation where the equilibrium number of signatories and the policy regime employed in the treaty is endogenously derived, it turns out that in most cases the quota regime is the preferred instrument by the majority of coalition members (FINUS/RUNDSHAGEN 1997).

Roughly speaking, the reason for the last two results is that an emission quota generates a rather even distribution of net benefits. This is particularly true if the base year from which emission reductions must be conducted is given by the Nash equilibrium of the stage game (see FINUS/RUNDSHAGEN 1997). This is what we will assume in the following.

Denote the fraction by which emissions must be reduced by R. Hence, emissions after abatement has been conducted are given by $E_i^Q = (1-R)E_i^N$, where Q stands for quota agreement and $0 \leq R \leq 1$. Using this information, a country's representatives can compute what percentage R would be optimal for their country. They will base their computation on the net benefit function of their country (1). Country i's proposal is given by

$$R_i = \frac{n\alpha_i(nb + n\alpha_i - b - \alpha)}{b^2 + 2b\alpha - 2nb\alpha_i + (\alpha - n\alpha_i)^2 + n^2 b\alpha_i} \quad \text{with} \quad \left.\frac{\partial R_i}{\partial \alpha_i}\right|_{\alpha=const.} > 0 \tag{10}$$

It is easily checked that $R_i \leq 1$ holds, but to ensure $0 \leq R_i$ we have to require $C_2 := (n-1)\, b \geq \alpha - n\alpha_i$. We assume C_2 to hold in the following in order to rule out corner solutions.

From (10) it is evident that the proposals will differ if environmental awareness varies across countries. A country's proposal is positively related to the parameter α_i. Hence, the country with the lowest value of α_i will put forward the lowest suggestion R_i. Denote this country by $i = 1$ for reference reason and thus $R_1 < R_i \; \forall \, i \neq 1$.

Of course, there are many possibilities to solve this bargaining problem. For instance one could choose a bargaining concept of cooperative game theory like the Nash bargaining solution or the Shapley value. This solution concepts are well founded in the game theoretical literature and possess some interesting axiomatic properties, but also satisfy some normative criteria. Among other factors, this explains there frequent application of these solution concepts in the environmental economics literature. (see, e.g., BOTTEON/CARRARO 1997 and BARRETT 1997).

In international politics, however, a different simple bargaining rule can frequently be observed: Bargaining partners agree on the lowest common denominator, henceforth abbreviated LCD-decision rule (ENDRES 1996, 1997a). Typical examples comprise the decision process within the European Union or within the UNO on particular issues (LENSCHOW 1996). The reason for this decision rule is simple: no country can be forced into an agreement, the accession must be voluntary if it happens at all. Besides this obvious reason the LCD-decision rule possess the interesting property that «truthtelling» is a dominant strategy. In other words, it does not pay a country to make a strategically manipulated offer R_i' instead of its «true» R_i in order to force the final agreement in its favor. Particularly, in environments with incomplete information such a property reduces the transaction costs of bargaining.

To demonstrate this result we only have to show that no country can do better than to stick to its proposal R_i. For the «bottleneck» country 1 – the country of which proposal is final accepted – this is obvious. Any proposal R_1' below ($R_1 < R_1'$) or above ($R_1' > R_1$) country 1's true proposal leads to a lower welfare since if R_1 is accepted, country 1 is in its «private equilibrium». For a country $i \neq 1$ a strategically biased proposal is only effective if it is below country 1's proposal R_1, $R_i' < R_1$. However, if such a proposal $R_i' < R_1$ is accepted, this implies a lower welfare for country i due to the assumption of a concave net benefit function and $E_i^Q \leq E_i^N$.

In contrast, if for instance the arithmetic mean of all suggestions is taken to be the abatement target laid down in an agreement, then neither an equilibrium in dominant strategies nor a Nash equilibrium exists in the bargaining game. Countries which like to see an emission reduction above the mean would submit a upward biased proposal. Those countries with a low interest in emission reductions would bias its proposal downward. Consequently, in the case of asymmetric countries ($\alpha_i \neq \alpha_j$), proposals would be permanently changed as a reaction to counter proposals with no convergence to a final agreement.

Therefore, in the following we assume that countries agree on the lowest common denominator. Due to our notational assumption, this is country 1' s proposal. Simple computations reveal that each country benefits from such an agreement compared to the Nash equilibrium of the stage game, irrespectively of the parameter values. Thus, a basic prerequisite of stability is satisfied. However, before investigating stability in a dynamic context, we briefly discuss the effects of an increase in environmental awareness on the incentive to defect.

3.2. *Stability Analysis in a Static Context*

Again, assuming that a country uses its best reply function when taking a free-ride, we can compute the incentive to free-ride $I_i = NB_i^D(E_i(E_{-i}^Q), E_{-i}^Q) - NB_i^Q(E_i^Q, E_{-i}^Q)$.

$$I_i^Q > 0 \;\forall\, i; \frac{\partial I_1^Q}{\partial \alpha_1} > 0; \frac{\partial I_1^Q}{\partial \alpha_i} < 0; \frac{\partial I_i^Q}{\partial \alpha_1} > 0; \frac{\partial I_i^Q}{\partial \alpha_i} < 0; \frac{\partial I_i^Q}{\partial \alpha_j} < 0 \;\forall\, i \neq j \neq 1 \tag{11}$$

Due to the assumption that country 1 is the bottleneck and therefore $\alpha_1 < \alpha_i \ \forall \ i \neq 1$ holds, the results presented in [11] only hold for marginal changes of environmental awareness. That is, the results only apply as long as country 1 remains the bottleneck country in the «process» of changing α_i. Otherwise, according to the LCD assumption, a different proposal would be accepted in the agreement and effects can not be signed anymore.

From (11) it is evident that raising environmental awareness in the bottleneck country increases the free-rider incentive in all countries including country 1. The opposite holds if environmental awareness increases in some other countries. In the latter case the gap between Nash equilibrium emissions and the quota solution narrows which reduces the incentive to defect. In the former case this effect is also at work, but is outweighed by the effect more ambitious abatement targets have on the incentive to free-ride $[\partial R_1/\partial \alpha_1 > 0$, see also (10)].

Thus, the «counter-intuitive» results derived in Section 2 still hold in an agreement on uniform emission reductions, though some signs of the derivatives have changed. This is confirmed when looking at the dynamics.

3.3. Stability Analysis in a Dynamic Context

Again, we assume that the trigger strategy described above is used to enforce the agreement. Then, the following relations can be established:

$$\frac{\partial \delta_1^{Min}}{\partial \alpha_1} < 0; \frac{\partial \delta_1^{Min}}{\partial \alpha_i} > 0; \frac{\partial \delta_i^{Min}}{\partial \alpha_1} > 0; \frac{\partial \delta_i^{Min}}{\partial \alpha_i} < 0; \frac{\partial \delta_i^{Min}}{\partial \alpha_j} \lessgtr 0 \ \forall \ i \neq j \neq 1 \tag{12}$$

Once more, an increase in environmental awareness in the bottleneck country increases the discount factor requirements in the neighbor countries and vice versa. Thus if a country discount factor is close to δ_i^{Min} an increase in environmental awareness may jeopardize stability of an agreement. This result is a paradox: On the one hand, one would argue that environmental awareness should be particularly raised in the bottleneck country because then, more ambitious abatement targets can be implemented. On the other hand, the results in (12) suggest that this makes it more difficult to control the other countries. Depending on the actual discount factors of countries, this second effect might contradict the «good intentions».

Again, there are two basic effects at work which are responsible for this result: the incentive effect and the punishment effect or deterrence effect (see, e.g., Section 2.3). However, in those cases in which they work in opposite directions the punishment effect is stronger. E.g., if environmental consciousness is raised in country 1 the incentive to defect in this country increases. But so do emissions in the Nash equilibrium in the other countries $[\partial E_i^N/\partial \alpha_1 > 0$, see (2)] and therefore a more severe punishment is at the disposal of these countries to control country 1. Obviously the latter effect dominates over the former. In contrast, if environmental awareness increases in country 1, the incentive to defect in a country i ($i \neq 1$) increases, but aggregate punishment emissions in all countries except i are reduced $(\partial(E_1^N + \sum E_{j \neq i}^N)/\partial\alpha_1 < 0)$. Hence, punishment becomes less deterrent for country i. In this case, both effects work in the same direction and reinforce each other.

Taken together, the results indicate that in an international policy setting with strategically behaving agents, many effects which work through different channels have to be taken into consideration when predicting an outcome. Intuitive reasoning is not sufficient and misguiding. Obviously, the common hold believe that increasing environmental awareness in a society will have only positive externalities on the stability of agreements has to be qualified.

4. THE ROLE OF UNILATERAL ACTIONS IN AN AGREEMENT ON A UNIFORM EMISSION REDUCTION QUOTA

In this section we investigate how unilateral actions taken by a country affect equilibrium emissions in an agreement on a uniform emission reduction quota. The analysis is motivated by the fact that sometimes environmental pressure groups demand such unilateral steps, intending to give a «good example» to other countries and thereby improving the environment (see HOEL 1991). In particular, we want to analyze what happens if either a country overfulfills the agreement *after* it has been signed (cases 1.1, 1.2 and 1.3) or if it unilaterally abates more *before* negotiations commence (case 2).

In order to keep matters simple and brief, we assume symmetric countries. This allows to derive some results in a straightforward manner because for this assumption it is true that a quota solution is socially optimal.

We refer to the country which takes unilateral actions as country 1 and for all other countries we use the index j. We investigate first the possibility that the public of country 1 demands from its government to overfulfill the agreement *after* it has been signed. Three subcases, denoted case 1.1, 1.2 and 1.3, can be distinguished, depending at which stage the commitment of overfulfillment is known to the parties involved and how they react to this knowledge.

Case 1.1:

One possibility is that the demand for overfulfillment is raised *after* the treaty has been signed. This case is particularly simple to analyze. In the first place, emission reductions are conducted according to (10) where $\alpha_1 = \alpha_j$ = a holds. Then, after country 1 has conducted the additional emission reduction, emissions are given by

$$E_{1(1.1)}^{Q} = \frac{bd}{b+n^2a} \cdot (1-k), \; E_{j(1.1)}^{Q} = \frac{bd}{b+n^2a} \; \forall \, j \neq 1, \tag{13}$$

where the subscript in brackets refers to the case, and k denotes the «portion of overfulfillment», $0 \leq k \leq 1$. Thus, of course, $E^Q_{1(1.1)} + (n-1) E^Q_{j(1.1)} < n E^Q_j$ for any $k > 0$, where we refer to E^Q_j as emissions without unilateral action. In this case the unilateral action has a «positive» effect on environmental quality because strategic reactions by neighbors are absent.

Since $\partial NB_i / \partial E_j < 0 \; \forall \; i$ and $E^Q_{j(1.1)} = E^Q_j$, a country $j \neq 1$, must benefit from the unilateral action of country 1, i.e., $NB^Q_{j(1.1)} - NB^Q_j > 0$. Since without unilateral action the quota solution is socially optimal for symmetric countries ($a_1 = \alpha_j = a$), and therefore constitutes a Pareto-Optimum by definition as well, country 1 must have suffered a welfare loss through its activity.

Case 1.2:

A second possibility to capture the effect of an overfulfillment (*after* the agreement has been signed) is to assume that the demand for unilateral action is raised in country 1 *before* negotiations take place. If this is known to countries $j \neq 1$, and if they believe that country 1 will respond to this demand, they will take this commitment into account when putting forward their proposal.

Moreover, if the government in country 1 knows that it will have to respond to this demand for overfulfillment it may consider this in its proposal as well.

Formally, in case 1.2 we assume that countries 1 and j assume $E^Q_1 = (1-k) \cdot E^N_1 \cdot (1-R)$ and $E^Q_j = (1-R) \cdot E^N_j$ when maximizing net benefits according to (1) in the text. Then, the proposals are given by

$$R_{1(1.2)} = \frac{bk^2 - na + n^2a + ak^2 - bk - nka}{b - 2bk + bk^2 + n^2a - 2nka + ak^2}, \; R_{j(1.2)} = \frac{a(k^2 - 2nk + n^2 - n)}{b + n^2a - 2nka + ak^2}. \tag{14}$$

To rule out corner solutions, we have to require C_3: $= k < n - \sqrt{n}$ to ensure $R_{j(1.2)} > 0$. Of course, C_3 is only binding in the case of $n = 2$ because of $0 \leq k \leq 1$. For any $n > 2$, C_3 is not binding. A sufficient condition to ensure $R_{1(1.2)} > 0$ is C_4: $= a > b/4$. Assume C_4 to hold. Then, $R_{1(1.2)} > R_{j(1.2)}$ and the uniform proposal of the j countries ($j \neq 1$), $R_{j(1.2)}$, is accepted. From this we can compute equilibrium emissions to be

$$E^Q_{1(1.2)} = \frac{bd(1-k)}{b + n^2a - 2nka + ak^2}, \; E^Q_{j(1.2)} = \frac{bd}{b + n^2a - 2nka + ak^2},$$
$$E^Q_{(1.2)} = \frac{bd(n-k)}{b + n^2a - 2nka + ak^2}. \tag{15}$$

Because the anticipated unilateral action leads to strategically motivated changes in the proposals and therefore in the reduction levels, the ecological and welfare effects in this case cannot be induced by simple reasoning. A comparison regarding aggregate emissions reveals

$$E^{Q} - E^{Q}_{(1.2)} = \frac{dbk\left(b - n^{2}a + nka\right)}{\left(b + n^{2}a\right)\cdot\left(b + n^{2}a - 2nka + ak^{2}\right)} \tag{16}$$

where E^{Q} ($E^{Q}_{(1.2)}$) denotes aggregate emissions without (with) unilateral action. The term in (16) can only be positive if $n = 2$. Due to C_4, $E^{Q} - E^{Q}_{(1.2)} < 0$ holds for $n > 2$. This implies that because countries $j \neq 1$ anticipate the unilateral action after the agreement, they «adjust» their proposal downward, which leads to lower abatement targets specified in the treaty. Obviously, though country 1 takes unilateral actions after the agreement has been signed, this «adjustment effect» is so strong that in most cases the ecological net effect is negative. Only for $n = 2$ and large k, it is possible that the unilateral action is not compensated by the adjustment effect.

Computing net benefits of a country $j \neq 1$, without (NB^{Q}_{j}) and with unilateral action ($NB^{Q}_{j(1.2)}$) gives

$$NB^{Q}_{j} = \frac{b^{2}d^{2}}{2\left(b + n^{2}a\right)}, \quad NB^{Q}_{j(1.2)} = \frac{b^{2}d^{2}}{2\left(b + n^{2}a - 2nka + ak^{2}\right)}. \tag{17}$$

It is straightforward to show that $NB^{Q}_{j(1.2)} > NB^{Q}_{j}$ for any $k > 0$. Again from this it follows that country 1's welfare must have decreased because the quota solution is a Pareto-Optimum in the symmetric country case if no unilateral actions are taken.

Case 1.3:

The third possibility assumes that government 1 is not influenced by the demand for unilateral actions in formulating its proposal. Only the other governments anticipate in their proposals the possible unilateral action of country 1 after the treaty has been signed.

Accordingly, country 1 assumes $E^{Q}_{1} = (1 - R)\cdot E^{N}_{1}$ and $E^{Q}_{j} = (1 - R)\cdot E^{N}_{j}$ when calculating its proposal $R_{j(1.3)}$. Thus, $R_{j(1.3)} = R_{i}$ with R_i as given in (10) in the text. In contrast, countries j assume $E^{Q}_{1} = (1 - R)\cdot E^{N}_{1}\cdot(1 - k)$ and $E^{Q}_{j} = (1 - R)\cdot E^{N}_{j}$ when putting forward their proposals. Hence, countries' j proposal is given by $R_{j(1.3)} = R_{j(1.2)}$, where $R_{j(1.2)}$ has been calculated in (14). A comparison shows that $R_{1(1.3)} > R_{j(1.3)}$ is true, implying that countries j are the bottleneck countries. But then equilibrium emissions and welfare levels are the same as derived in case 1.2 and what has been discussed there applies.[11]

11. A similar result can be obtained in the opposite case where only country 1 adjusts its proposal and the other $j \neq 1$ countries do not use the information of the unilateral action strategically.

Case 2:

Finally, we consider the possibility that a country unilaterally reduces emissions *before* negotiations take place. This implies that country *1* has to reduce its emissions from a lower base level compared to the Nash equilibrium when an emission reduction target *R* is agreed upon. If country *1* reduces its emission by the portion *k* before negotiations start, then its base year emissions are given by $E_1^{N*} = (1-k) \cdot E_1^N$, where E_1^N denotes the initial emission level and E_1^{N*} the emission level after the action has been taken. But then a country *1*'s proposal assumes $E_1^Q = (1-k) \cdot E_1^N \cdot (1-R)$ and $E_j^Q = (1-R) \cdot E_j^N$, and a country *j*'s proposal is based on this assumption as well. This is, however, the underlying assumptions of case 1.2 and hence, again, the results derived in that case apply.

Taken together, the results strongly indicate that unilateral actions open up various opportunities to countries to strategically adjust their proposals. This is particularly true if unilateral actions are taken before an agreement is signed or the unilateral action after an agreement is anticipated. This may lead to results not intended by those who advocated unilateral actions. Whereas in HOEL (1991) a similar result is due to the fact that unilateral actions weaken the bargaining position of countries, in our bargaining setting this «negative» result comes about because countries reduce their abatement efforts when observing or anticipating unilateral actions by others.

Admittedly, the results are derived from a stylized model. In particular, we assumed identical countries. Nevertheless, the message is clear: Sometimes good intentions are not enough, in fact, they may even worsen the situation. This problem is particularly virulent in an international setting in which governments behave strategically when pursuing their immediate interests.[12] Moreover, the results suggest that whether an overfulfillment by a country has a positive effect crucial depends upon when the commitment is known.

5. SUMMARY AND CONCLUSIONS

The role of environmental awareness to solve the problems associated with the protection of the global commons has been analyzed in an economic framework. We started out by analyzing a socially optimal agreement. It became apparent that raising environmental awareness in a country may very well increase the stability requirements for some other countries. This was demonstrated in a static and a dynamic context. It was shown that for some countries a socially optimal agreement might imply a lower welfare than the Nash equilibrium of the stage game. This is particularly true if environmental consciousness varies across different societies. Since socially optimal conventions are hard to find in real world we extended our analysis beyond this «idealistic» framework. In particular,

12. This is an argument similar to the one presented in STEPHAN/IMBODEN (1995), pp. 206.

we considered a more pragmatic approach to reduce global externalities – a uniform emission reduction quota.

We assumed that if countries hold different views regarding this quota, they settle for the lowest common denominator (LCD). Such a solution generates relatively equal net benefits to each country (compared to alternative solution concepts) and ensures that each country is better off than in the Nash equilibrium. Moreover, it was shown that the bargaining outcome is an equilibrium in dominant strategies. This makes the LCD-decision rule «immune» against strategically manipulated proposals which helps to explain its frequent appearance in international politics, even though it usually leads to inefficient outcomes.

However, also in a quota agreement the ambiguity of environmental awareness on the stability of a treaty was demonstrated. Paradoxically, raising environmental awareness in the bottleneck country, thereby favoring the approval of more demanding abatement targets, leads to higher stability requirements in all other countries. Also, a change in environmental consciousness in other countries might very well make it more difficult to control some signatories within an agreement.

Finally, we demonstrated within a simple bargaining setting on emission reduction quotas that unilateral actions taken by a country before or after an agreement has been signed may not deliver the results intended in the first place. In fact, only if these actions are taken after an agreement has been signed and are not anticipated by neighbor states, the ecological effect will be positive. In most other cases strategic behavior of countries' representatives leads to just the opposite effect of what advocates of unilateral actions have in mind. The strategic behavior of countries lowers the abatement target agreed upon in the negotiations. This more than compensates the positive effect of unilateral actions. Thus, unilateral actions do not only influence the bargaining position of countries unfavorably as has been suggested by MÄLER (1990) and demonstrated by HOEL (1991), but also make countries more reluctant to agree on demanding abatement targets because they expect that others will do the job.

Of course, the results have been derived from a stylized model and therefore should be taken with «a grain of salt» before being digested. In particular, we assumed that only net benefits from emissions enter into governments' objective function. Assuming that politicians pursue also other goals like increasing their status or the probability of being re-elected might change some of the results.[13] Nevertheless, the results indicate that good intentions are sometimes not enough and may have effects not intended initially. Particularly in international pollution control, where no institution exists which can enforce agreements, policy recommendations have to take strategic behavior of agents into consideration and should not rely on intuitive reasoning.

13. It should be kept in mind that aggregate net welfare maximizing behavior of politicians is not necessarily identical to vote maximizing behavior, due to distributional issues.

REFERENCES

BARRETT, S. (1992). International Environmental Agreements as Games. In: Pethig, R. (ed.), Conflicts and Cooperation in Managing Environmental Resources. *Microeconomic Studies*, Berlin et al.: Springer, chapter 1, 11–37.

BARRET S. (1994). Self-Enforcing International Environmental Agreements. «*Oxford Economic Papers*», vol. 46, pp. 878–894.

BARRETT, S. (1997). Heterogenous International Agreements. In: Carraro, C. (ed.), International Environmental Negotiations: Strategic Policy Issues. Edward Elgar, Cheltenham. pp. 9–25.

BERNHOLZ, P. and F. BREYER (1994). Grundlagen der Politischen Ökonomie. Band 2: Ökonomische Theorie der Politik. Tübingen.

BOTHE, M. (1995). The Evaluation of Enforcement Mechanisms in International Law. In: Wolfrum, R. (ed.), Enforcing Environmental Standards: Economic Mechanisms as Viable Means? Berlin et al.: Springer, pp. 13–38.

BOTTEON, M. and C. CARRARO, (1997). Burden-Sharing and Coalition Stability in Environmental Negotiations with Asymmetric Countries. In: Carraro, C. (ed.), International Environmental Negotiations: *Strategic Policy Issues*. Cheltenham: Edward Elgar, pp. 26–55.

CARRARO, C. and D. SINISCALCO, (1992). The International Dimension of Environmental Policy. «*European Economic Review*», vol. 36, pp. 379–387.

DESVOUGES, W.H., V.K. SMITH and H.H. RINK, (1989). Communicating Radon Risks Effectively. Washington D.C. (U.S. EPA).

ENDRES, A. (1996). Designing a Greenhouse Treaty: Some Economic Problems. In: Eide, E. and R. Van Den Bergh (eds.), *Law and Economics of the Environment*. Oslo, pp. 201–224.

ENDRES, A. (1997a). Negotiating a Climate Convention – The Role of Prices and Quantities, «*International Review of Law and Economics*», vol. 17, pp. 147–156.

ENDRES, A. (1997b). Increasing Environmental Awareness to Protect the Global Commons – A Curmudgeon's View. «Kyklos», vol. 50, S. 3–27.

ENDRES, A. and M. FINUS (1996). Umweltpolitische Zielbestimmung im Spannungsfeld gesellschaftlicher Interessengruppen: Theorie und Empirie. In: Siebert, H. (ed.), Elemente einer rationalen Umweltpolitik. Expertisen zur umweltpolitischen Neuorientierung. Tübingen: J.C.B. Mohr, pp. 35–133.

ENDRES, A. and M. FINUS (1998). Renegotiation-Proof Equilibria in a Bargaining Game over Global Emission Reductions – Does the Instrumental Framework Matter? Forthcoming in Hanley, N. und H. Folmer (eds.), Game Theory and the Global Environment. Cheltenham: Edward Elgar.

ENDRES, A. and K. HOLM-MÜLLER (1998). Bewertung von Umweltschäden – Theorie und Praxis sozioökonomischer Verfahren. Stuttgart: Kohlhammer.

FINUS, M. (1992). Der Reisekostenansatz – Darstellung, methodische Probleme und Anwendungsmöglichkeiten. Gießen.

FINUS, M. and B. RUNDSHAGEN (1996). Renegotiation-Proof Equilibria in a Global Emission Game when Players are Impatient. Discussion paper No. 236, Hagen: University of Hagen, forthcoming in Journal of Environmental and Resource Economics.

FINUS, M. and B. RUNDSHAGEN (1997). Toward a Positive Theory of Coalition Formation and Endogenous Instrumental Choice in Global Pollution Control. Discussion paper No. 239, Hagen: University of Hagen, forthcoming in Public Choice.

FREEMAN, M.A. (1993). The Measurement of Environmental and Resource Values, Washington D.C. (Ressources for the Future).

HANLEY, N. and C.L. SPASH (1993), Cost Benfit Analysis and the Environment. Aldershot/Hants – Brookfield/Vermont.

HANLEY, N., J.F. SHOGREN and B. WHITE (1997). Environmental Economics, London: MacMillan.

HERBER, B.P. and J.T. RAGA (1995). An International Carbon Tax to Combat Global Warming, «*The American Journal of Economics and Sociology*», vol. 54, pp. 257–276.

HOEL, M. (1991). Global Environmental Problems: The Effects of Unilateral Actions Taken by One Country. «*Journal of Environmental Economics and Management*», vol. 20, pp. 55–70.

HOEL, M. (1992). International Environment Conventions: The Case of Uniform Reductions of Emissions. «*Environmental and Resource Economics*», vol. 2, pp. 141–159.

HOWARTH, R.B. and R.B. NORGAARD (1992). Environmental Valuation under Sustainable Development. «*American Economic Review*», vol. 80, pp. 473–477.

JAEGER, F. (1994). Natur und Wirtschaft, Chur: Rüegger.

LENSCHOW, A. (1996). Der umweltpolitische Entscheidungsprozeß in der Europäischen Union am Beispiel der Klimapolitik. In: Brauch, H.G. (Hrsg.), Klimapolitik. Berlin et al.: Springer, chapter 7, pp. 89–104.

MÄLER, K.-G. (1990). International Environmental Problems. «Oxford Review of Economic Policy», vol. 6, pp. 80–108.

MOHR, E. (1994). Environmental Norms, Society and Economics. «*Ecological Economics*», vol. 9, pp. 229–239.

MÜLLER-FÜRSTENBERGER, G. und G. STEPHAN (1997). Environmental Policy and Cooperation Beyond the Nation State: An Introduction and Overview. Forthcoming in «*Structural Change and Economic Dynamics*», vol. 8, pp. 99–114.

SCHELLING, T.C. (1960). The Strategy of Conflict. Cambridge: Harvard University Press.

STEPHAN, G., G. MÜLLER-FÜRSTENBERGER and P. PREVIDOLI, (1997). Overlapping Generations or Infinitely-Lived Agents: Intergenerational Altruism and the Economics of Global Warming. «*Environmental and Resource Economics*», vol. 10, pp. 27–40.

STEPHAN, G. and D. IMBODEN (1995). Laissez-faire, Kooperation oder Alleingang: Klimapolitik in der Schweiz. «Schweizerische Zeitschrift für Volkswirtschaft und Statistik», vol. 131, pp. 203–226.

SUMMARY

It is commonly believed that an increase in environmental awareness is conducive to the protection of the global commons. Moreover, it has been proposed that governments should «give a good example» by unilaterally reducing emissions. In a simple analytical framework it is shown that these relations are less straightforward than intuitively expected. On the contrary: «Good intentions» might lead to lower abatement targets within an international environmental agreement and to lower stability due to strategic behavior of governments.

ZUSAMMENFASSUNG

In der Öffentlichkeit wird häufig argumentiert, dass ein höheres Umweltbewusstsein der Bevölkerung, der Bewahrung globaler Umweltgüter zuträglich wäre. Ebenfalls wird oftmals gefordert, Staaten sollten vor oder nach dem Abschluß eines internationalen Umweltschutzabkommens unilaterale Umweltschutzmassnahmen durchführen, um anderen Ländern ein gutes Beispiel zu geben. Von einem solchen Verhalten werden insgesamt positive Effekte für die Umwelt erwartet. Mit einem einfachen Modell wird gezeigt, daß die Zusammenhänge nicht so eindeutig sind wie zunächst vermutet. Im Gegenteil: Oftmals führen «gute Absichten» aufgrund des strategischen Verhaltens von Regierungen zu niedrigeren Emissionsreduktionszielen innerhalb eines Umweltschutzabkommens sowie zu einer geringeren Vertragsstabilität.

RESUME

Il est souvent soutenu qu'une prise de conscience approfondie de l'importance de l'environnement de la part de la population est propice à la conservation de biens environnementaux globaux. Il est aussi fréquemment postulé que les états devraient enforcer des mesures unilaterales de protection de l'environnement avant la conclusion d'un accord international de protection de l'environnement afin de donner le bon exemple aux autres pays. Un tel comportement devrait géneralement produire des effets positifs pour l'environnement. La présente contribution démontre à l'aide d'un modèle simple que ces relations ne sont pas aussi évidentes qu'elles ne le paraissent à première vue. Bien au contraire: du à un comportement stratégique des gouvernements, les «bonnes intentions» mènent souvent à des buts de réduction d'émissions inférieurs au

sein d'un accord de protection de l'environnement ainsi qu'à une moindre stabilité du traite.