Environmental and Resource Economics **9:** 153–170, 1997.

Incentives to Participate in an International Environmental Agreement

MICHAEL HOEL[1] and KERSTIN SCHNEIDER[2,*]
[1] *Department of Economics, University of Oslo, P.O. Box 1095 Blindern, N-0317 Oslo;*
[2] *Department of Economics, University of Dortmund, D-44221 Dortmund, Germany*
(* *corresponding author*)

Accepted 18 June 1996

Abstract. For international environmental problems involving many countries, such as, e.g., the climate problem, it is unlikely that all countries will participate in an international environmental agreement. If some countries commit themselves to cooperate, while the remaining countries act independently and in pure self-interest, it appears to be possible to achieve a Pareto improvement if the non-signatory countries reduce their emissions, in exchange for transfers from the countries which sign an agreement. However, the paper shows that the prospect of receiving a transfer for reducing one's emissions provided the country does not commit itself to cooperation, tends to reduce the incentive a country might have to commit itself to cooperation. Moreover, if the disincentive effect of such side payments is strong, total emissions will be higher in a situation with side payments than in a situation in which the signatory countries commit themselves to not give transfers to free riding countries.

Key words: international environmental agreement, cooperation, side payments

1. Introduction

Consider an environmental problem for which the quality of the environment in each country depends not only on its own emissions, but also on the emissions of other countries. Obvious examples of such international environmental problems are global warming, the depletion of the ozone layer, and acid rain, but there are also numerous other examples. For such problems, it is well known that one needs some kind of international environmental agreement (IEA) in order to achieve a socially optimal outcome. However, cooperation in international environmental issues is not easy to achieve, since each country may have an incentive to be a free rider. On the one hand, each country may be better off participating in an agreement than it would be without any agreement. On the other hand, however, there appear to be incentives for countries to stay outside the agreement and to simply pursue their self-interest, i.e. to be a free rider. For international environmental problems involving many countries, such as the climate problem, it therefore seems unlikely that all countries will participate in an international environmental agreement.

A situation in which some countries commit themselves to cooperate, while the remaining countries act independently and in pure self-interest, will generally be socially inefficient. Starting with such a situation, it could be possible to achieve a Pareto improvement if the non-cooperating countries reduce their emissions in exchange for transfers from the countries which are committed to cooperate. Such a mechanism for expanding the number of countries who reduce their emissions below the 'business as usual' level is often referred to as 'joint implementation'; see, e.g., Barrett (1993a,b), Bohm (1994), Hanish et al. (1992), and Johnson (1993a,b).

Given the group of countries which is committed to cooperate, it is in principle possible to achieve a Pareto improvement with an appropriately designed system of joint implementation.[1] However, one could argue that the prospect of receiving a transfer for reducing one's emissions will reduce the incentive a country might have to sign an agreement.[2] Even if some kind of mechanism for joint implementation gives a Pareto improvement over the case without any side payments for a *given number of committed countries*, ruling out any kind of joint implementation might therefore nevertheless be best, since the number of countries committed to cooperate may be higher in the latter case. The issue of whether a system of side payments reduces incentives to cooperate or not is the issue of the present paper. In order to address this issue, we first discuss why some countries commit themselves to cooperation in the first place.

Why should some countries commit themselves to cooperation? A coalition of countries which commit themselves to cooperate is stable if it is not in the self-interest of any of these countries to break out of the coalition. The size of such a stable coalition depends on what happens if a country chooses to be a free rider instead of cooperating. In Carraro and Siniscalco (1993), Barrett (1994a), and Hoel (1992), for example, it is assumed that all countries first decide (independently and simultaneously) whether or not to cooperate, after which the cooperating and non-cooperating countries decide upon their emission levels. In a game of this type, the only negative consequence for a country of not joining the coalition is that the optimal emission levels of the remaining cooperating countries may increase. Only if this increase in emission levels hurts the defector more than the costs it saves by defecting, will it be optimal for a potential defector to cooperate. The studies mentioned above argue that for problems such as the climate problem the number of countries in a stable coalition is likely to be very small. Moreover, total emissions from all countries will not be much lower than they are in the non-cooperative equilibrium.

The assumptions that the decisions of whether or not to cooperate are made simultaneously and once and for all are not particularly realistic. Bauer (1993) has described a situation in which the decision of one country of whether or not to cooperate may affect the corresponding decisions of other countries. In this model there exist equilibria with many countries cooperating. A similar approach is to extend the two stage games of the type described above to a repeated game. It is well known from the literature on game theory that it may be possible to sustain tacit cooperation as a perfect equilibrium of a non-cooperative (infinitely) repeated game; see, e.g., Torvanger (1993) for a discussion in the context of international environmental agreements. The fact that decisions about greenhouse gas emissions are frequently repeated may thus solve the free rider problem. However, as repeated games of this type have multiple equilibria, the coordination problems of reaching a Pareto optimal equilibrium are large. Obviously, these coordination problems are larger the larger the number of countries involved. It therefore seems likely that only a subset of all countries will commit themselves to cooperation. Barrett (1994a) has also shown that if one restricts oneself to those equilibria of the repeated game which are renegotiation proof, the number of cooperating countries may be quite small.

The literature discussed above assumes that each country is only concerned with its own welfare level, defined (loosely) as income minus environmental costs. However, countries typically incur additional, non-environmental costs if they choose not to join the environmental agreement. Social norms and conventions, for example, may play an important role in sustaining international environmental agreements. At the individual level, there is a lot of behavior which is easiest explained by social norms and conventions; see, e.g., Elster (1989). This may very well apply at the country level as well: A government may feel uncomfortable if it breaks the social norm of sticking to an agreement of reduced emissions, even if in strict economic terms it may benefit from being a free rider. Alternatively, a country could be excluded from trade agreements which imposes costs on the defecting country. The non-environmental cost is likely to be higher the larger the number of cooperating countries. In the present paper, a social norm of the above type or other non-environmental international links between countries plays an important role in the determination of how many countries commit themselves to cooperate.

In Section 2, we present the general framework and give a precise definition of a stable coalition of cooperating countries. Sections 3 and 4 derive the stability conditions that apply to the particular model used in the analysis without and with side payments, and Section 5 compares the incentives to cooperate under the two cases. A numerical example is given in Section 6, and some conclusions are drawn in Section 7.

2. Stable Coalitions

Let $u_i^C(n)$ denote the utility of country i if it is a member of the cooperating coalition, which we denote by $I(n)$, when there are n countries in the coalition (including country i). The utility of country i if it is outside the coalition $I(n)$ is denoted by $u_i^F(n)$.

A self-enforcing coalition $I(n)$ of n countries satisfies the following conditions:

$$u_i^C(n) \geqslant u_i^F(n-1) \quad i \in I(n) \tag{1}$$

and

$$u_i^F(n) \geqslant u_i^C(n+1) \quad i \notin I(n). \tag{2}$$

Equations (1) and (2) state that for a coalition to be stable, it must neither pay for a coalition member to defect nor for a country outside the coalition to join the agreement. Note that u_i depends only on the number of countries in the coalition, but due to the symmetry of the countries it does not matter which particular country breaks the agreement.

In the next two sections the stability conditions will be derived for agreements with and without side payments to the non-cooperating countries. Throughout, we make the following assumptions.

All countries have the same production function, where output $r(x)$ is a function only of emissions x (i.e., other inputs are held constant). It is assumed that $r' > 0$ and $r'' < 0$.

Environmental costs are equal for all countries and are given as a linear function of total emissions. Marginal environmental costs of each country are denoted by b.

If country i chooses not to cooperate when n other countries are cooperating, it incurs non-environmental costs $\alpha_i(n)$.[3] The cost functions $\alpha_i(n)$ may differ across countries, and for all countries it is assumed that this cost is increasing in n. For small values of n it is not unreasonable to assume $\alpha_i(n) = 0$, at least for some countries. Countries are indexed so that $\alpha_i(n) \geqslant \alpha_2(n) \geqslant \ldots \geqslant \alpha_N(n)$ for all n.[4]

It is assumed that the cooperating countries and the non-cooperating countries choose their emission levels simultaneously. However, due to the assumption of linear environmental costs, the results would be the same if we instead had assumed that the non-cooperating countries chose their emission levels after having observed the emission levels of the cooperating countries.

3. Cooperation Without Side Payments

When there are no side payments, and n countries are cooperating, they choose a common emission level in order to maximize the utility level of each member. We thus have

$$u_i(n) = \max_x\{r(x) - b[nx + (N - n)y]\}, \tag{3}$$

where N is the total number of countries sharing the environmental resource. Total emissions are equal to $b[nx + (N-n)y]$. The n countries in the coalition each emit x, while the $N - n$ countries outside the coalition each emit y. From (3) we get the first order condition for a cooperating country

$$r'(x) - bn = 0. \tag{4}$$

Note that due to the constancy of marginal environmental damage, the level of emissions chosen by the cooperating countries does not depend on y, the emissions of non-cooperating countries. It is easily seen from (4) that emissions x of each cooperating country are a declining function of the marginal damage, b, and the number of countries in the coalition, n. Hence joining countries are rewarded by reduced emissions of the members of the coalition. Defecting countries are punished through increased pollution by the cooperating countries. Furthermore, in the absence of other policy instruments, the level of abatement acts as the only punishment mechanism available to the coalition.

Each of the non-cooperating countries can only control their own emission level, and therefore regard emissions from all other countries as given. The utility level of country i if it does not cooperate and n other countries have committed themselves to cooperation is therefore given by

$$u_i^F(n) = \max_y[r(y) - by] - b[nx + (N - n - 1)y^0] - \alpha_i(n), \tag{5}$$

where y^0 is the emission level of each of the $(N - n - 1)$ other cooperating countries. The term $\alpha_i(n)$ is the non-environmental cost for country i of breaking the agreement, and it is assumed that this cost term may differ across countries. Hence $\alpha_i(n)$ is a function accounting for heterogeneity across countries. It helps to explain why countries join an agreement and choose not to be a free rider, even without an international legal framework that can force countries to cut back on pollution.

Taking x and y^0 as given, the maximization problem defined by (5) has the following first-order condition

$$r'(x) - b = 0. \tag{6}$$

The emission level given by (6), which we have denoted by y^0, thus makes the marginal product of emissions in each non-cooperating country equal to the marginal damage caused by the polluting activities in the country itself. As before, the level of emissions y^0 does not depend on the emissions from cooperating countries, because the marginal environmental damage is constant. Moreover, since $n \geqslant 1$ it follows from (4) and (6) that $y^0 \geqslant x$.

The resulting utility of the cooperating and non-cooperating countries, respectively, is

$$u^C(n) = r(x(n)) - b[nx(n) + (N - n)y^0] \tag{7}$$

and

$$u_i^F(n) = r(y^0) - b[nx(n) + (N - n)y^0] - \alpha_i(n). \tag{8}$$

Note that the utility of countries in the coalition is the same for all countries $i = 1, \ldots, n$, so that we can omit the subscript i. However, utility of a free riding country i depends also on country specific non-environmental costs of breaking the agreement.

Define $\Phi_i(n, \alpha_i(n)) \equiv u^C(n) - u_i^F(n - 1)$, i.e., from (7) and (8)

$$\Phi_i = r(x(n)) - r(y^0) + b[-nx(n) + (n-1)x(n-1) + y^0] + \alpha_i(n-1). \tag{9}$$

The definition of a stable coalition (given by (1) and (2)) may now be written as

$$\begin{aligned}\Phi_i(n) &\geqslant 0 \;\; for \;\; i \in I(n)\\ \Phi_i(n) &\leqslant 0 \;\; for \;\; i \notin I(n).\end{aligned} \tag{10}$$

Hence, for country n to sign the agreement, output gains from free riding must be offset by increased environmental costs due to free riding and the increased emissions of the remaining cooperating countries, augmented by $\alpha_i(n)$.

From (9) and our assumption that countries are indexed so that $\alpha_1(n) \geqslant \alpha_2(n) \geqslant \ldots \geqslant \alpha_N(n)$, we know that $\Phi_i(n)$ is increasing in its index i. The largest possible stable coalition must therefore be the one consisting of the first n countries, so that (10) in this case may be rewritten as

$$\begin{aligned}\Phi_n(n) &\geqslant 0\\ \Phi_{n+1}(n+1) &\leqslant 0.\end{aligned} \tag{11}$$

The size of the largest stable coalition depends on the properties of the function $\Phi_i(n)$, which may be written as

$$\Phi_i(n) = \phi(n) + \alpha_i(n-1), \tag{12}$$

where

$$\phi(n) = [r(x(n)) - bnx(n)] - [r(y^0) - by^0] + b(n-1)x(n-1). \tag{13}$$

In Appendix A, it is shown that $\phi(2) > 0$ and that $\phi(n)$ is decreasing in n for $n > 0$.[5] However, $\alpha_i(n)$ is increasing in n, so we do not know in which direction $\Phi_i(n)$ changes as n increases. If $\Phi_i(n)$ is decreasing in n (for given i), then it must be true that $\Phi_n(n)$ is decreasing in n. Generally, however, $\Phi_n(n)$ may decrease or rise with n.

Figures 1 and 2 illustrate two possible cases. In Figure 1, $\Phi_n(n)$ is declining in n, and eventually becomes negative. In this case n^* is the largest stable coalition.[6] In Figure 2, the largest stable coalition is N, i.e., the coalition of all countries. Notice that n^* is in this case also a stable coalition.

4. Cooperation With Side Payments

As mentioned in the Introduction, it might be in the interest of countries who have committed themselves to cooperate to induce the remaining countries to reduce

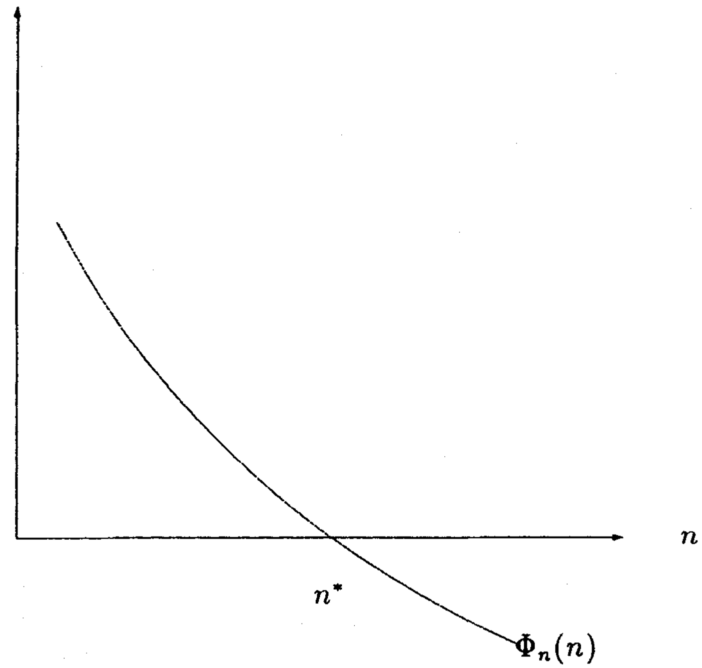

Figure 1. Largest stable coalition n^*, no side payments.

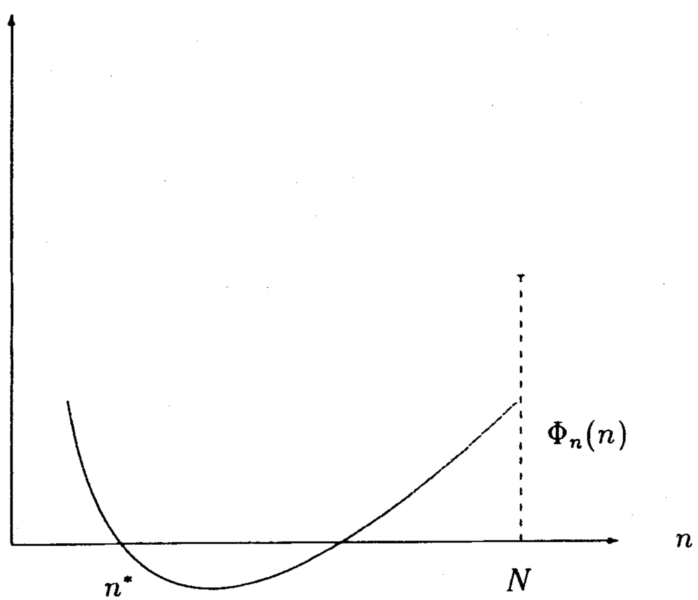

Figure 2. Full cooperation, no side payments.

their emissions by means of side payments. Clearly, the reason is that pollution is not constraint to national boundaries. The original group of cooperating countries, henceforth called the signatories, benefit from the reduction of emissions by non-signatories. Hence signatory countries might be better off if they compensate non-signatory countries for reducing environmental damage.

For this case with side payments, the signatories choose not only their own emission level, but also the emission level of the non-signatories. They must also take into consideration how much they must pay the non-signatories in order to induce them to reduce their emissions to the specified level. Formally the utility level of each of the signatories is now given by

$$\begin{aligned} u^C(n) &= \max_{x,y} n[r(x) - b[nx + (N-n)y]] - \\ &\quad (N-n)[r(y^0) - r(y) - b(y^0 - y)]. \end{aligned} \tag{14}$$

As before, x is the emission level of the signatories, while y is the emission level of the non-signatories. In the absence of side payments y^0 is the free-rider emission level. The second term of (14) thus represents payments necessary to reduce the non-signatories' emissions to y. Clearly, they have to be reimbursed for output losses due to reduced emissions, minus the benefits they get from reduced pollution. In other words, side payments have to be tailored such that the utility level of each of the non-signatory countries remains unchanged, before and after the transfer.

Note that in this simple model side payments are paid to every non-signatory country, while the political discussion typically demands to pay low-income regions for reduced pollution. Indirectly this argument is included in the α_i-function. $\alpha_i(n)$ captures the non-environmental costs of not joining an agreement. It could be argued that low-income countries are those who have very little to lose if they do not sign an IEA. Thus they are not likely to participate and are the recipients of side payments.

The first order conditions for the signatory countries are

$$r'(x) = bn \tag{15a}$$

$$r'(y) = b(n+1) \tag{15b}$$

from which it follows that $y(n) = x(n+1)$.

x as well as y are decreasing in n, the number of countries signing the agreement. Furthermore, (15a) and (15b) show that signatories want non-signatories to emit like a member of a coalition of size $n+1$.

As before, the coalition will be stable if conditions (1) and (2) hold. This implies that the expressions for the utility levels to be compared are

$$\begin{aligned} u^C(n) &= r(x(n)) - b[nx(n) + (N-n)y(n)] \\ &- \frac{N-n}{n}[r(y^0) - r(y(n)) - b(y^0 - y(n)] \end{aligned} \tag{16}$$

and

$$\begin{aligned} u_i^F(n-1) &= r(y(n-1)) - b[(n-1)x(n-1) \\ &+ (N-n+1)y(n-1)] + [r(y^0) - r(y(n-1)) \\ &- b[y^0 - y(n-1)] - \alpha_i(n-1). \end{aligned} \tag{17}$$

Define $\Psi_i(n, \alpha_i(n)) \equiv u_i^C(n) - u_i^F(n-1) \geqslant 0$.

From (16) and (17) we get

$$\begin{aligned} \Psi_i(n, \alpha_i(n)) &= -b[nx(n) - (n-1)x(n-1) + (N-n)y(n) \\ &- (N-n+1)y(n-1)] \\ &- \frac{N}{n}[r(y^0) - by^0] + \frac{N-n}{n}[r(y(n)) - by(n)] \\ &+ r(y(n-1)) - by(n-1) + \alpha_i(n-1, N) \geqslant 0. \end{aligned} \tag{18}$$

The definition of a stable condition is given by (10) as before, except that now Ψ takes the place of Φ. Similarly, the largest stable coalition is given by (11) with Φ replaced by Ψ.

5. Gains From Cooperation With Versus Without Side Payments

Having obtained expressions for the gains from cooperation, equations (11) and (18) can be compared, giving an answer to the question whether side payments reduce or increase incentives to join a coalition.

In this comparison, we shall assume that the α_i-functions are the same in the case with side payments as without. This is by no means obvious: One could argue that if one institutionalizes a system of side payments to non-signatories in order to induce them to reduce their emission levels, it becomes more acceptable to be a non-signatory. If this is the case, the α_i-functions ought to be lower-valued in the case with side payments than in the case without. This factor would unambiguously tend to make signing an agreement less attractive in the case with side payments than without. Our assumption about equal α_i-functions means that we are comparing the incentives to sign the agreement in the two cases when we ignore this possible difference.

Subtracting Ψ_i (given by (18)) from Φ_i (given by (11) or (12)–(13)) and rearranging, we find

$$\Pi(n) \equiv \Phi_i(n) - \Psi_i(n) = \frac{N-n}{n}\pi(n), \tag{19}$$

where

$$\pi(n) = [r(y^0) - by^0] - [r(y(n)) - by(n)] + nb[y(n) - y(n-1)]. \tag{20}$$

The two first terms taken together are clearly positive since y^0 maximizes $r(y) - by$. These two terms give a country's utility reduction of reducing its own emissions from y^0 to $y(n)$, and will be larger the larger n is (since $y'(n) < 0$). The last term is negative and measures the increase in environmental costs a country suffers if n countries increase their emissions from $y(n)$ to $y(n-1)$. For large n, we expect the first two terms to dominate the last term, so that the sign of $\pi(n)$ will be positive. In Appendix B, we give a more detailed argument for why $\pi(n)$ will be positive for sufficiently large values of n, and have also proved that $\pi(n) > 0$ for $n > 2$ for the case of $r''' = 0$.

Provided $\pi(n)$ is positive for the relevant n-values, the curve for $\Psi_n(n)$ will lie below the curve for $\Phi_n(n)$ for $n < N$, as illustrated in Figures 3 and 4 for the two cases previously illustrated in Figures 1 and 2. Notice however that it follows from (19) that we always must have $\Psi_N(N) = \Phi_N(N)$. This means that if the coalition of all countries is stable in the case without side payments, the same coalition is stable in the case with the side payments. If the maximal stable coalition n^* in the case without side payments is smaller than N (as in Figure 3), then the maximal stable coalition n^{**} in the case with side payments is smaller than n^* (see Figure 3).

The comparison of gains from signing an agreement with and without side payments strongly suggests that fewer countries tend to join an agreement if side payments are allowed for. This result raises another question. Is total pollution lower in a world with or without side payments?

Denote the size of the group of signatories under the alternative regimes as n^* and n^{**}, where the superscripts * and ** refer to cases without and with side payments, respectively. Total emissions for the two cases are equal to

$$n^*x(n^*) + (N - n^*)y^0$$

and

$$n^{**}x(n^{**}) + (N - n^{**})y(n^{**}).$$

Clearly, two opposing effects are at work. From what was said earlier, the number of countries joining the coalition decreases if side payments are feasible. This

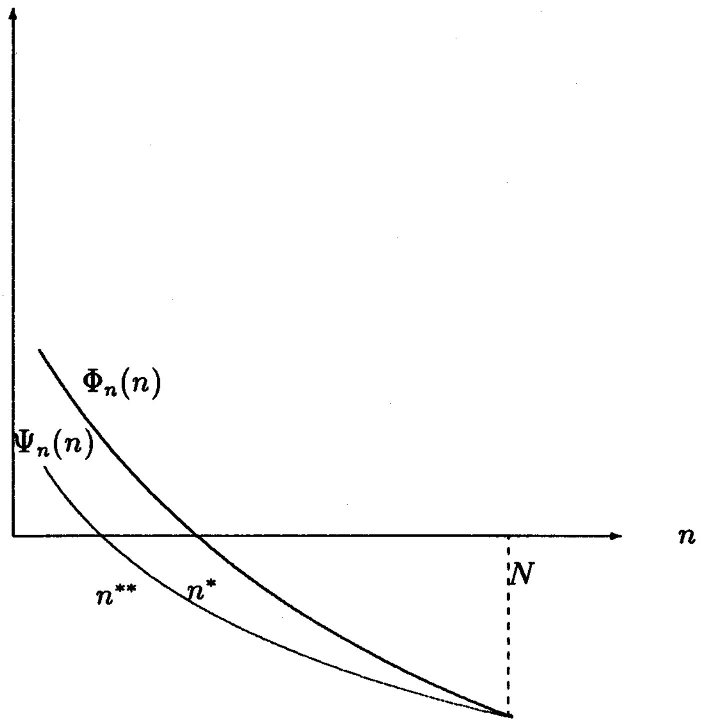

Figure 3. Largest stable coalition with and without side payments.

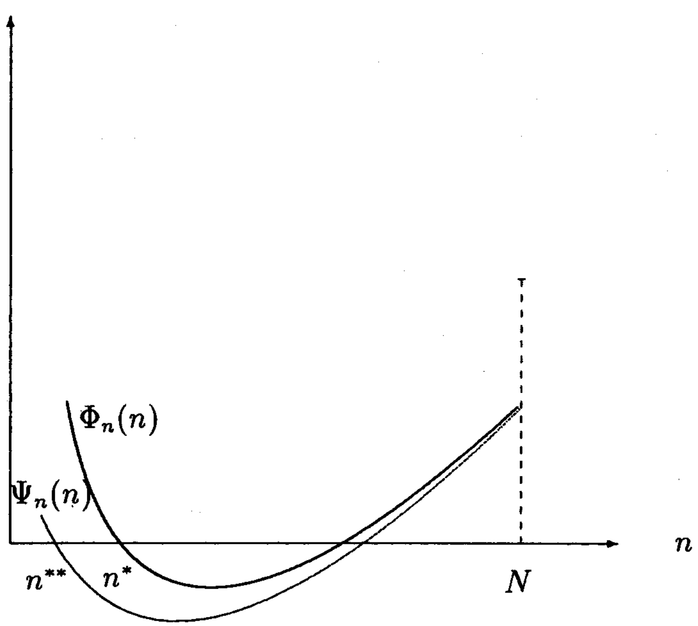

Figure 4. Full cooperation with and without side payments.

reduces the number of countries emitting $x(n)$. Since emissions are decreasing in the number of countries in the coalition, $x(n^{**}) > x(n^{*})$. Each country in the coalition emits more; on the other hand, the pollution of the non-signatory countries is reduced when they receive side payments as compared to the free-riding level, hence $y(n^{**}) < y^0$. The overall effect on emissions cannot be determined unambiguously at this level of generality. To get a better understanding of the workings of the model, an example is employed in the following section.

6. An Example

To illustrate the theoretical model, consider a production technology of the following kind:

$$r(x) = x - 0.5x^2. \tag{21}$$

Environmental costs for each country are

$$\frac{\beta}{N}(nx + (N - n)y). \tag{22}$$

Note that $b = \beta/N$, i.e., marginal environmental costs depend on the number of countries sharing the environmental resource. To interpret β, consider the first-best social optimum, i.e., the emission level maximizing $r(x) - bNx = r(x) - \beta x$. This is given by $r'(x) = \beta$, or, given (21), $x = 1 - \beta$. We may therefore interpret β as the optimal relative reduction in x due to environmental costs. It follows that the typical range for β is [0, 1]. Also, the highest environmental cost a country could face is β, if each country sets $x = 1$, irrespective of environmental damage incurred. Notice also that with our specification of environmental costs, the optimal relative abatement is equal to β, independent of how many countries the world is divided into.

Given the specification (21)–(22) it follows from (19) that

$$\Pi = \frac{1}{2}\frac{\beta^2}{N^2}(n - 2)(N - n). \tag{23}$$

The incentives to join the coalition are thus stronger if no side payments are allowed for, given that more than two countries form a coalition.[7] The difference grows as β, the socially optimal reduction in emissions, gets larger.

To solve for the number of cooperating countries, the non-environmental cost of breaking the agreement α_i needs to be specified. To keep things simple, let the α_i-function be the same for all countries and given by

$$\alpha(n) = \frac{1}{4}\beta^2\frac{n-1}{N-2}.$$

Using this α-function ensures that the non-environmental cost is zero if no countries form a coalition ($n = 1$). If all but one countries join the coalition ($n = N - 1$), the cost of being outside the coalition is high for the free-riding country, i.e., $0.25\beta^2$, which amounts to almost half of the output loss of the free-riding country if it had chosen to sign the agreement.

The number of countries in the coalition in the absence of side payments is

$$n^* = \begin{cases} N & \text{for } N \leqslant 8 \\ 3 + \frac{N^2}{2(N-2)} & \text{for } N > 8 \end{cases}$$

and

$$n^{**} = \begin{cases} N & \text{for } N \leqslant 8 \\ 1 + \frac{4-6N+2N^2}{8-8N+N^2} & \text{for } N > 8 \end{cases}$$

if we allow for side payments.[8]

The resulting number of cooperating countries for different values of N, the total number of countries, is given in Table I. As expected, the number of countries in the coalition is always at least as large in a system without side payments as in a system allowing for side payments. The more countries share the resource, the greater is the difference in incentives to join an agreement. While it is possible to achieve full cooperation, full cooperation can only be achieved if the number of countries is really small. As N gets larger, n^{**} is equal to three, whereas n^* is close to $N/2$. Hence, at least in this example, cooperation exists and stable coalitions result with and without side payments. Moreover, the number of signatory countries when side payments are not allowed for greatly exceeds the coalition size when side payments are part of the agreement.

This raises another issue: The effectiveness of international agreements. Tables II(a) and II(b) report the total emissions, given n^* and n^{**} for different values of N. In the example considered, total emissions are generally lower, when side payments are not part of an IEA, and the total number of countries is large. That result is not too surprising, because the difference between the number of countries joining the coalition with and without side payments is increasing in N.

Clearly, this discussion is merely based on an example. Policy recommendations cannot be derived from the results presented without caution. However, the intuition that side payments granted in IEAs reduce the incentive to join an agreement has been confirmed. Furthermore, side payments might not even be desirable from an environmental point of view, because they might decrease the degree of cooperation sufficiently to result in higher total emissions.

7. Conclusions

When global environmental problems are at stake, no country can be forced to adhere to an internationally announced level of abatement. Only voluntary partic-

Table I. Number of countries in a self-enforcing coalition.

N	No side payments	Side payments
5	5	5
6	6	6
7	7	7
8	8	8
9	8	7
10	9	6
20	14	3
30	19	3
40	24	3
50	29	3
60	34	3
70	39	3
80	44	3
90	49	3
100	54	3

Table 2a. Total emissions: no side payments.

N	β			
	0.1	0.3	0.6	0.9
8	7.2	5.6	3.2	0.8
10	9.18	7.54	5.08	2.62
20	18.99	16.97	13.94	10.91
30	28.76	26.28	22.56	18.84
40	38.52	35.56	31.12	26.68
100	97.04	91.11	82.23	73.34

Table 2b. Total emissions: side payments.

N	β			
	0.1	0.3	0.6	0.9
8	7.2	5.6	3.2	0.8
10	9.36	8.08	6.16	4.24
20	19.62	18.85	17.69	16.54
30	29.61	28.83	27.66	26.49
40	39.61	38.82	37.65	36.47
100	99.6	96.81	97.62	96.43

ipation in an agreement is possible. Thus, when the international community discusses potential strategies to reduce CO_2 emissions this aspect should not remain unnoticed.

Various papers have shown that the number of countries participating in an agreement is likely to be small. The present paper confirms this only partly. Accounting for country-specific non-environmental costs increases the number of participating countries without side payments substantially. In the example of Section 6 the participation rate is about one half. While non-environmental costs also affect the decision of a country when side payments are part of the agreement, the influence seems diminished. Allowing for side payments reduces the number of participating countries in the example to only three.

Furthermore, a second question arises: Is it desirable, from an environmental point of view, to pay countries outside the coalition to reduce emissions? It turns out that if the disincentive effect of side payments is strong, such that only a small fraction of all countries joins a coalition, total emissions with side payments are higher than without.

Clearly, the paper leaves questions that ought to be addressed at some point. The game modelled is a one-shot game. A natural extension to the analysis is to look at a repeated game. Intuitively one would expect the degree of cooperation to increase under both regimes. It seems unlikely though that the main results of the one-shot game will be reversed. However, a formal analysis is needed to fully understand the effects of repeated games on incentives to cooperate in a self-enforcing environmental agreement.

Acknowledgement

Financial support from the Nordic Council of Ministers is gratefully acknowledged.

Notes

1 As the literature above demonstrates, however, there may be severe practical problems.

2 Countries are called cooperating or signatory countries if they reduce emissions below the free riding level *without* requiring compensation. Countries that do not reduce emissions without being compensated are non-signatory or non-cooperating countries, even though cutting back on emissions in return for compensation can be interpreted as some degree of cooperation or commitment.

3 One problem with interpreting $\alpha_i(n)$ as costs arising from trade sanctions is that trade sanctions are likely to affect the cooperating countries as well (Barrett 1994b). The model presented here does not capture these interaction effects.

4 We are thus assuming that the ranking of countries by their cost functions α_i is independent of the argument value n in $\alpha_i(n)$.

5 To be precise: A sufficient condition for $\phi'(n) > 0$ for $n > 2$ is that $r''' \geqslant 0$.

6 More precisely, the largest stable coalition is the largest integer which does not exceed $n*$.

7 This is a consequence of (19), which implies $r''' = 0$. See Appendix B.

8 See note 6.

Appendix A: Properties of $\phi(n)$

From (13) we have

$$\phi(2) = [r(x(2)) - 2bx(2)] - [r(y^0) - by^0] + bx(1). \tag{A1}$$

From (4) and (6) we know that $y^0 = x(1)$, hence

$$\phi(2) = [r(x(2)) - 2bx(2)] - [r(x(1)) - 2bx(1)]. \tag{A2}$$

Since $x(2)$ maximizes $r(x) - 2bx$, it follows that $\phi(2)$ is positive.

To see how $\phi(n)$ depends on n, we rewrite (13) as

$$\phi(n) = \max_x[r(x) - nbx] - \max_y[r(y) - by] + (n-1)bx(n-1). \tag{A3}$$

Differentiating and using the envelope theorem yields

$$\phi'(n) = -bx(n) + bx(n-1) + (n-1)bx'(n-1). \tag{A4}$$

Moreover, from the mean value theorem we have

$$x(n) = x(n-1) + x'(n-\delta), \tag{A5}$$

where $\delta \in [0,1]$, which inserted in (A4) gives

$$\phi'(n) = -bx'(n-\delta) + (n-1)bx'(n-1),$$

or

$$\phi'(n) = (n-2)bx'(n-1) + b[x'(n-1) - x'(n-\delta)]. \tag{A6}$$

We know from (4) that $x'(n) = b/r''(x(n)) < 0$, so that the term in square brackets is non-positive if $r''' \geqslant 0$. Since $(n-2)bx' < 0$ for $n > 2$, we thus have $\phi'(n) < 0$ for $n > 2$ if $r''' \geqslant 0$.

Appendix B: Properties of $\pi(n)$

Equation (20) may be rewritten as

$$\pi(n) = [r(y^0) - by^0] - [r(y(n)) - (1+n)by(n)] - bny(n-1)$$

or, since $y(n)$ maximizes $(r(y) - (1+n)by)$ (cf. (15b))

$$\pi(n) = \nu(0) - \nu(n) - bny(n-1), \tag{B1}$$

where

$$\nu(n) = \max_y[r(y) - (1+n)by]. \tag{B2}$$

From the mean value theorem we have

$$\nu(n) = \nu(0) + n\nu'(m), \tag{B3}$$

where m is some number in the interval $[0,n]$. Moreover, applying the envelope theorem to (B1) yields

$$\nu'(m) = -by(m). \tag{B4}$$

Inserting (B3) and (B4) into (B1) gives us

$$\pi(n) = nb[y(m) - y(n-1)]. \tag{B5}$$

It is thus clear that $\pi(n)$ is positive, provided $y(m) > y(n-1)$, i.e., provided $m < n-1$. For n sufficiently large, it is likely that the appropriate m in (B2) will be smaller than $n-1$.

For the case in which $r''' = 0$, it is useful to use the second order Taylor expansion instead of (B2), which gives

$$\nu(n) = \nu(0) + n\nu'(0) + \frac{1}{2}n^2\nu'', \tag{B6}$$

where

$$\nu'' = -by' = -\frac{b^2}{r''} \tag{B7}$$

from (B4) and (15b). If ν'' had depended on n, (B6) would only hold for ν'' evaluated at a particular argument value (in the interval $(0,n)$). However, in the present case it follows from (B7) that ν'' is constant since r'' is constant. Inserting (B6) into (B1), and using (B4) and (B7), gives

$$\pi(n) = nby^0 - nby(n-1) + \frac{n^2b^2}{2r''}. \tag{B8}$$

Since y' is constant ($= b/r''$) and $y(0) = y^0$ (from (6) and (15b)) we can write

$$y(n-1) = y^0 + (n-1)y' = y^0 + (n-1)\frac{b}{r''}. \tag{B9}$$

Inserting (B8) into (B7) and rearranging gives us

$$\pi(n) = \frac{nb^2}{r''}\left(1 - \frac{n}{2}\right). \tag{B10}$$

Since $r'' < 0$, it follows that $\pi(n) > 0$ for $n > 2$.

References

Barrett, S. (1993a), 'Joint Implementation for Achieving National Abatement Commitments in the Framework Convention on Climate Change', revised draft, London Business School (for OECD).

Barrett, S. (1993b), 'A Strategic Analysis of "Joint Implementation" Mechanisms in the Framework Convention on Climate Change', revised draft, London Business School (for UNCTAD), May 26.

Barrett, S. (1994a), 'Self-Enforcing International Environmental Agreements', *Oxford Economic Papers* **46** (October, special issue on environmental economics), 878–894.

Barrett, S. (1994b), 'Trade Restrictions in International Environmental Agreements', CSERGE working paper: GEC; 94, 13.

Bauer, A. (1993), 'International Cooperation over Environmental Goods', Mimeo, University of Munich.

Bohm, P. (1994), 'On the Feasibility of Joint Implementation of Carbon Emissions Reductions', Research papers in economics, Dept. of Economics, University of Stockholm.

Carraro, C. and D. Siniscalco (1993), 'Strategies for the International Protection of the Environment', *Journal of Public Economics* **52**, 309–328.

Elster, J. (1989), 'Social Norms and Economic Theory', *Journal of Economic Perspectives* **3**, 99–117.

Hanish, T., R. K. Pachauri, D. Schmitt, and P. Vellinga (1992), 'The Climate Convention: Criteria and Guidelines for Joint Implementation', Policy Note 1992: 2, CICERO, Oslo.

Hoel, M. (1992), 'International Environmental Conventions: The Case of Uniform Reductions of Emissions', *Environmental and Resource Economics* **2**, 141–159.

Johnson, I. (1993a), 'Operational Criteria for Joint Implementation', OECD/IEA, International Conference on the Economics of Climate Change, Paris, June 14–16.

Johnson, I. (1993b), 'Operational criteria for Joint Implementation; A Response', OECD/IEA, International Conference on the Economics of Climate Change, Paris, June 14–16.

Torvanger, A. (1993), 'Efficient Contracts in a Game of Nations Pursuing Greenhouse Gas Emissions Abatement', CICERO Working Paper 93: 3, Center for Climate and Energy Research, Oslo.