Minimum participation rules and the effectiveness of multilateral environmental agreements

Author(s):
Rutz, Samuel

Publication date:
2001

Permanent link:
https://doi.org/10.3929/ethz-a-004296584

Originally published in:
Working paper series / WIF 2001(22)

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

WIF
Institut für Wirtschaftsforschung
Center for Economic Research

Working Paper Series

01/22

Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agreements

Samuel Rutz

July 2001

Center for Economic Research
Swiss Federal Institute of Technology
Weinbergstrasse 35
CH-8006 Zurich

Telefon: 0041-1-632 41 28
Telefax: 0041-1-632 12 18
email: library@wif.gess.ethz.ch

Minimum Participation Rules and the Effectiveness of Multilateral Environmental Agreements [1]

Samuel Rutz

Center for Economic Research

Swiss Federal Institute of Technology

CH-8952 Zurich

Tel ++41 1 632 50 04

Fax ++41 1 632 10 42

e-mail rutz@wif.gess.ethz.ch

ABSTRACT: The goal of this paper is to assess the effects of minimum participation rules on the outcome of multilateral environmental negotiations. Although such rules are widely used in practice, up to now, minimum participation levels have hardly been taken into consideration when modelling negotiations on transboundary pollution problems. By incorporating a minimum participation rule in a widely used game-theoretic model of coalition formation, we show that such a "hand-tying" mechanism has the potential to overcome, at least partially, the freerider incentives inherent in transboundary pollution abatement. Besides these theoretical implications, the new model allows to formulate a simple hypothesis about how many countries will ratify a multilateral environmental agreement in practice. We will study the relevance of this hypothesis by analyzing the accession data of 122 existing agreements. Our analysis reveals that in most cases more than the predicted number of countries ratify a treaty, which allows us to draw conclusions about the effectiveness of existing multinational environmental agreements.

July, 2001

[1] The authors are grateful to Thomas Borek, Martin Brown, Matthias Gysler, Armin Schmutzler and Renate Schubert for helpful comments and suggestions.

1 Introduction

Every multilateral environmental agreement is based on a treaty text which is subject to ratification by its parties. Such treaty texts specify in the first place the measures to be taken by the parties to deal with the environmental problem at hand. A treaty text settles however also procedural questions, like e.g. accessions, withdrawals, amendments, reservations, entry into force etc. From an economic point of view, the article about "entry into force" is of particular interest, since in most cases it contains a minimum participation rule. A minimum participation rule specifies the number of instruments of ratification, acceptance, approval or accession that have to be deposited, before a treaty enters into force.[2]

Up to now, minimum participation rules have hardly been taken into consideration when modelling multilateral environmental negotiations.[3] A minimum participation level can however crucially influence the outcome of negotiations, since, from a game-theoretical point of view, it potentially changes the structure of the game. While in the literature it is commonly assumed that multilateral environmental negotiations are social dilemma games, the introduction of a minimum participation rule has the potential to overcome, at least partially, the freerider incentive present in this class of games.

In this paper we will incorporate a minimum participation rule in a widely used and often cited two-stage game of multilateral environmental negotiations. Results of this "standard" game suggest that, due to the prevailing freerider incentive, no efficient cooperation between countries can be expected. The introduction of a minimum participation rule however changes the decision structure of the standard game to a game of chicken. In this class of games there are generally many non-cooperative equilibria, each of which may be optimal and none of which is dominant, and the task of the players is to coordinate their actions to select one.

[2] The words acceptance and approval basically mean the same as ratification, but are less formal and therefore sometimes preferred by states which have constitutional difficulties with the term ratification. Accession is the method used by a state to become party to a treaty which it did not sign whilst the treaty was open for signature. For the sake of briefness, we will make no difference between these instruments in the following, but sum all of them under the term "ratification".

[3] As far as we are aware, the only paper studying minimum participation levels is by Black et al. (1993). In contrast to the approach of this paper, they analyse the size of the optimal minimum participation level in an incomplete information framework.

Thus, contrary to social dilemma games, real efficiency gains are feasible in a game of chicken, given that the players efficiently solve the coordination problem. The necessary condition for the achievement of an effective agreement in our case is that a minimum participation rule is binding, i.e. the required number of ratifications must be set higher than the number of countries that would have acceded an agreement anyway.

The introduction of a minimum participation rule not only has theoretical implications, but allows to formulate a simple hypothesis about how many countries will ratify an agreement in practice. The analysis of the accession data of 122 multilateral environmental agreements shows however that in most cases the actual number of signatory countries is substantially larger than our prediction. Given the inherent freerider incentive in transboundary pollution problems, this fact leaves us with a theoretical puzzle.

As a suggestion for further research, we will propose two extensions of the discussed game, which possibly could explain this phenomenon. Both of them relate to the negotiation process between countries, which is not explicitly modelled in the considered game. This process logically takes place before countries decide whether or not to ratify a treaty text and determines the feasible gains from cooperation. If, for reasons to be discussed later on, these negotiations lead to terms of a treaty which are not much stricter than what a country would have done unilateral, it is the ineffectiveness of such agreements which suggests itself as an explanation for the observed accession behavior.

The claim that many multilateral environmental agreements tend to be ineffective in practice is not new in the literature (see e.g. Endres et al. 2000). Due to the problem that information about cost and benefits of environmental measures is hard to find in a transboundary context, empirical evidence on the subject is however sparse up to now. Although the simple test we conduct in this paper cannot be considered as empirical evidence, it seems to support this claim.

The structure of the paper is as follows: In section 2 we address the question whether minimum participation rules are relevant in practice. As we will see, such rules are indeed a widely used instrument in multilateral environmental negotiations. In section 3, we introduce the above mentioned standard game and discuss solution concept and results. In the following section we will then incorporate a minimum participation rule in the standard game and analyze its effect on the outcome of negotiations. Section 5 derives a simple hypothesis from the equilibrium predictions of this new game. The relevance of this hypothesis will then be studied by analyzing accession data of 122 multilateral environmental agreements in section 6. The results of this analysis will leave us with a theoretical puzzle, which will be the subject of section 7. Finally, section 8 contains a summary and some concluding remarks.

2 The Importance of Minimum Participation Rules

Before analyzing the theoretical effects of minimum participation rules in more detail, let us address the question of the relevance of such rules in practice. To this purpose we studied the 122 treaty texts of multilateral environmental agreements and protocols provided by the *Center for International Earth Science Information Network* (CIESIN).[4]

Our study reveals, that minimum participation levels are a widely used instrument in practice: Only in 2 out of the 122 considered treaty texts no minimum participation clause was found. In 81 cases the participation rule merely asks for a certain number of ratifications, i.e. as soon as the specified number of countries has deposited an instrument of ratification the agreement enters into force. In another 22 cases, the minimum participation rule requires unanimous ratification by all of the initially negotiating parties. For the remaining 17 cases the participation rule is coupled to other requirements, i.e. for these agreements it is not sufficient that a certain number of countries ratify the treaty, but these countries additionally have to satisfy other criteria. The *Montreal Protocol on Substances that Deplete the Ozone Layer* may serve as an example. It asks for 11 instruments of ratification representing at least two-thirds of 1986 estimated global consumption of the controlled substances. These findings are summarized in table 1.[5]

[4] Center for International Earth Science Information Network (CIESIN). 1996-2001. Environmental Treaties and Resource Indicators (ENTRI) [online]. Palisades, NY: CIESIN. URL: http://sedac.ciesin.org/entri/.

[5] A detailed list of the studied treaties is available from the authors on request.

type of participation rule	number of treaties
simple participation rule	81
unanimity participation rule	22
participation rule & other requirements	17
no participation rule	2
total	**122**

Table 1

Table 1 clearly confirms the importance of minimum participation rules in practice. Therefore, it seems desirable to theoretically analyze the effects of such participation rules on the outcome of multilateral environmental negotiations.

In what follows, we will give a description of the structure and timing of the above mentioned standard model of multilateral environmental negotiations. This discussion will however be kept as short as possible, since the model is well-known in the literature. For a detailed discussion of the model we refer to Barrett (1994, 1997) or Rutz and Borek (2000). Further, we will introduce some general properties of the payoff function, which will proof helpful for the analysis in section 4, where a minimum participation rule is incorporated in the standard game.

3 The Standard Model

3.1 General Structure of the Model and Previous Literature

The model of multilateral environmental negotiations we analyze in this section is a static, complete information two-stage game: At the first stage countries decide non-cooperatively whether or not to sign an agreement. At the second stage both signatory and non-signatory countries set their abatement levels, where the signatory countries act as a single player and divide resulting gains from cooperation according to a given cooperative bargaining rule.

The decision of countries at the second stage can be modelled simultaneously or sequentially. In the sequential version of the game it is assumed that the signatory countries choose their abatement level first. Observing this, the non-signatory countries then choose their own. In the simultaneous version of the game signatory and non-signatory countries choose their abatement level at the same time.

Due to mathematical problems of solving the game explicitly for the optimal abatement levels and the size of the stable coalition, most results in the literature rest on numerical simulations. For the sequential version of the game, Barrett (1994, 1998) derived stable coalitions for the case of identical, as well as heterogeneous countries. For the simultaneous version Carraro and Siniscalco (1993) and Hoel (1994) derived stable coalitions for the case of identical countries and Botteon and Carraro (1998) for the case of heterogeneous countries.

Approximating the game by assuming a continuum of players, Rutz and Borek (2000) show however in a recent paper that an analytical solution to the two versions of the game can be found. Much of the discussion in section 3.2 and 3.3 is based on this work. Since results of the two versions do not differ substantially, we will concentrate on the sequential version of the game in what follows. Further, to keep matters simple, we assume that all countries are identical.

3.2 Solution and Results

The complete information two-stage game of multilateral environmental negotiations then has the following structure:[6]

> *Stage 1:*
>
> There are N identical countries, which decide simultaneously whether or not to sign an agreement.[7]
>
> *Stage 2:*
>
> First, the k signatory countries choose their aggregate abatement level $Q_s = \sum_{i=1}^{k} q_{s_i}$, where $q_{s_i} \geq 0$ denotes the individual abatement level of signatory country i, by maximizing their collective net benefits.

[6] In the following index s stands for signatory countries and index n for non-signatory countries.

[7] The assumption of identical countries makes it unnecessary to specify a burden-sharing rule at the first stage. This, because any of the burden-sharing rules derived from cooperative game theory leads in this case to an equal split of the cost of abatement, i.e. the burden-sharing rule does not matter when assuming identical countries.

The $(N-k)$ non-signatory countries then observe Q_s and simultaneously choose their individual abatement level $q_{n_j} \geq 0$ to maximize their individual net benefit.

The net benefit to non-signatory country j is given by the payoff function $\Pi_{n_j} = B(Q) - C(q_{n_j})$, where $Q = Q_s + Q_n$ and $Q_n = \sum_{j=k+1}^{N} q_{n_j}$, while the net benefit to a signatory country can be expressed as $\Pi_{s_i} = B(Q) - C(q_{s_i})$. Inherent to the transboundary pollution problem, the benefits of abatement $B(\cdot)$ typically depend not only on the individual adopted abatement level, but on the abatement undertaken by all countries. The costs of abatement $C(\cdot)$, however, are always incurred by the individual country. Standard assumptions in the literature are that the benefit function is increasing and concave ($B_Q > 0$, $B_{QQ} < 0$) and the cost function is increasing and convex ($C_q > 0$, $C_{qq} > 0$).

The game is then solved by backward induction. Note that the assumption of identical countries implies that all non-signatory countries will choose the same abatement level $q_n^*(k)$ and all signatory countries the same abatement level $q_s^*(k)$ in the equilibrium. A subgame perfect equilibrium further requires that all decisions made at the first stage, to join or not to join the coalition, are mutual best responses for all countries. This equilibrium condition is usually termed as the "stability condition". A coalition k is said to be "stable" if the following two conditions are met: First, there is no incentive for a signatory country to leave the coalition, i.e. no signatory country can increase its payoff by leaving the coalition. Second, there is no incentive for a non-signatory country to join the coalition, i.e. a non-signatory country cannot increase its payoff by joining the coalition. Formally these two conditions can be expressed as follows:

(1) $\Pi_s(k^*) \geq \Pi_n(k^*-1)$ and $\Pi_n(k^*) \geq \Pi_s(k^*+1)$,

where k^* denotes the size of the stable coalition.

As noted above, due to the fact that it is normally not possible to solve the game explicitly for $q_s^*(k^*)$ and $q_n^*(k^*)$, most results in the literature rest on numerical simulations. Stable coalitions for this sequential version of the game, using specific payoff functions, have been derived by Barrett (1994). His results suggest that the stable coalition usually involves only a small fraction of the negotiating countries and that gains from cooperation are very modest. Rutz and Borek (2000), working with the above specified general class of payoff functions, show however that these results are not surprising, since resulting efficiency gains are solely due to an "integer effect", i.e. coalition size being a discrete variable. Approximating the game by assuming a continuum of players, they show that the optimal choice of the abatement level for the signatory countries is the one they would have chosen in a situation without cooperation between countries, i.e. the non-cooperative abatement level. This then implies by definition of the reaction function that every non-signatory country will as well choose the non-cooperative abatement level.

To summarize, the results of the above model are not very encouraging, since they predict no efficient cooperation between countries or cooperation on a very low level, i.e. countries do not commit to any substantial improvement of environmental quality. In other words, the discussed coalition game cannot overcome the freerider incentives inherent in transboundary pollution abatement. However, the model neglects an important feature of real-life negotiations on transboundary pollution problems, namely minimum participation rules.

3.3 General Properties of the Payoff Function

As a preliminary step to facilitate the following analysis, it is worth to study some general properties of the payoff function of signatory and non-signatory countries in more detail. In particular, we are interested in the question how the payoff of the two types of countries changes when coalition size changes. This will help to derive an easy and intuitive solution to the game when considering minimum participation rules. We proceed by stating three propositions and then give a graphical illustration and an intuition of their content. Rigorous proofs of the three propositions can be found in Rutz and Borek (2000).

Proposition 1 *There exists exactly one intersection of the payoff function of the signatory countries, $\Pi_s(k)$, and the non-signatory countries, $\Pi_n(k)$, in the range $(0, N)$. Denote the coalition size associated with this intersection as k_0.*[8] *At k_0 the*

[8] *Note that k_0 needs not to be an integer number.*

following condition holds: $\Pi_s(k) = \Pi_n(k) = \Pi_{nc}$, *where* Π_{nc} *stands for the non-cooperative payoff, i.e. the payoff resulting in a situation without any cooperation between countries.*

Proposition 2 *For any coalition size smaller than* k_0, *the payoff to the non-signatory countries is smaller than the non-cooperative payoff* Π_{nc}: $\Pi_n(k) < \Pi_{nc}\ \forall k < k_0$.

Proposition 3 *For any other coalition size than* k_0, *the payoff to the signatory countries is bigger than the non-cooperative payoff* Π_{nc}: $\Pi_s(k) > \Pi_{nc}\ \forall k \neq k_0$. *That is,* k_0 *is a minimum of* $\Pi_s(k)$.

The three propositions together imply a relationship between the payoff of signatory and non-signatory countries and the coalition size as depicted in figure 1.

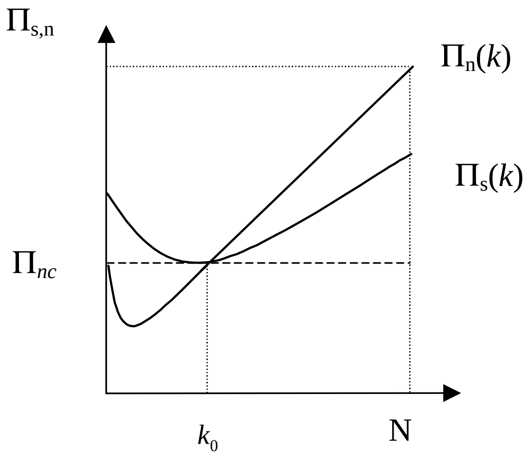

Figure 1: Payoff functions of the standard game

The intuition for the fact that the intersection of the payoff functions occurs exactly at the level of the non-cooperative payoff is the following: The sequential structure of the second stage of the game implies a "first-mover" advantage for the signatory countries. This "first-mover" advantage can however only effectively be exploited when the coalition of the signatory countries is small ($k < k_0$). In such a situation, the signatory countries take a freeride on the abatement effort of the non-signatory countries. This explains why for a small size of the coalition, the payoff of the signatory countries is bigger than the non-cooperative payoff, while the payoff of the non-signatory countries is strictly less than in a situation without cooperation. The bigger the coalition gets however, the weaker is the "first-mover" advantage, since there are simply less non-signatory countries left to freeride on. At k_0 the effect of the "first-mover" advantage completely disappears and for every coalition bigger than k_0 the situation is reversed. Now we are in the classical situation, where it benefits the non-signatory countries to take a freeride on the abatement efforts of the signatory countries.

Since it benefits the signatory countries to freeride when the coalition is smaller than k_0 and the non-signatory countries when the coalition is bigger than k_0, it is intuitively clear that the stability condition (Eq. (1)) can only be satisfied for a coalition size near k_0. To be more precise, a coalition smaller than k_0 can never be an equilibrium of the game, since the payoff to the non-signatory countries for any $k < k_0$ is strictly less than in a situation without cooperation between countries (see proposition 2). Therefore, any stable coalition must be bigger or equal to k_0. Since k_0 is not necessarily an integer number, the stable coalition will normally be larger than k_0. It will however not be much larger than k_0, since, in general, the freerider benefits of the non-signatory countries are increasing in the coalition size, which implies that the incentive to withdraw from the coalition gets stronger the bigger the coalition is. Thus, as the stable coalition k^* will not be far from k_0, only modest gains from cooperation can be expected in the equilibrium of the game.

Having understood these properties of the payoff function and their implications for the equilibrium of the game, it is now easy to see the effect of a minimum participation rule on the outcome of the game.

4 Effects of a Minimum Participation Rule

4.1 Structure of the Game with a Minimum Participation Rule

In what follows, we will describe the structure and the solution of the standard game when introducing a minimum participation rule. We will refer to this extension of the standard game as the "MPR-game". To keep matters simple, we will assume that the minimum participation level ($\overline{k}$) is given exogenous. This assumption is of course questionable, since the agreement on a certain number of countries is an integral part of the negotiations. Possible effects of a endogenous minimum participation rule will however be discussed in more detail in section 7.3. The complete information two-stage game we analyze in this section has the following structure:

Stage 1:

There are N identical countries, which decide simultaneously whether or not to sign an agreement, given an exogenous minimum participation level $\overline{k}$.

Stage 2:

- *if* $k < \overline{k}$*:* No agreement is reached and all N countries individually maximize their payoff function without regard to the choice of the other countries.
- *if* $k \geq \overline{k}$*:* The k signatory countries first choose their aggregate abatement level $Q_s = \sum_{i=1}^{k} q_{s_i}$ by maximizing their collective net benefits. The $(N-k)$ non-signatory countries then simultaneously choose their individual abatement level $q_{n_j} \geq 0$ to maximize their individual net benefit.

The introduction of a minimum participation level thus changes the game in the sense that for every coalition k smaller than $\overline{k}$ no agreement will be reached and consequently all countries will choose the non-cooperative abatement level q_{nc}. For every resulting coalition k bigger or equal to $\overline{k}$ however, the individual countries optimize as discussed in section 3.2. This implies that the payoff function of the signatory and the non-signatory countries, in dependence of the coalition size, no longer is continuous. This is illustrated in figure 2, where up to $\overline{k}$ all countries get the non-cooperative payoff. At $\overline{k}$ however, the payoff jumps back onto the "old payoff functions", i.e. on the payoff functions from the standard game.

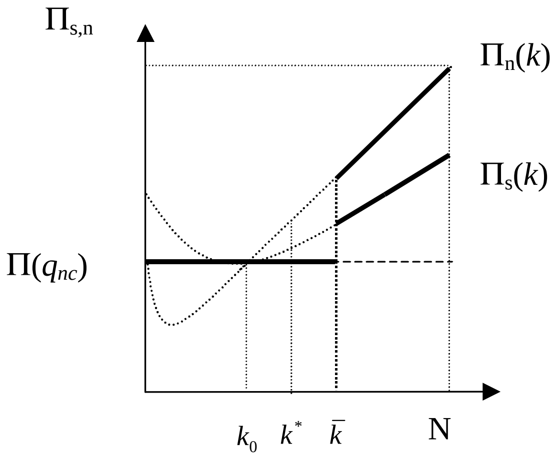

Figure 2: Payoff functions of the MPR-game

4.2 Equilibria of the MPR-game

To determine the equilibria of the MPR-game, we have to distinguish two cases: $\overline{k} < k_0$ and $\overline{k} \geq k_0$. In what follows, we will first check for these two cases if $\overline{k}$ is an equilibrium of the game and then look for other possible equilibria.

In the first case, when $\overline{k} < k_0$, it is immediately clear that $\overline{k}$ cannot be an equilibrium of the MPR-game. This is so, because for any $\overline{k} < k_0$ the right part of the stability condition (Eq. (1)) is not satisfied. This can easily be checked in the left picture of figure 3, but it follows as well straight from propositions 1-3, which tell us that for any $\overline{k}$ smaller than k_0, $\Pi_n(\overline{k}) < \Pi_{nc} < \Pi_s(\overline{k})$. In such a situation there is always an incentive for non-signatory countries to join the coalition, since this will increase their payoff. Therefore, when the minimum participation level is set too low, $\overline{k}$ will not be an equilibrium of the MPR-game. This result is summarized in the following proposition.

Proposition 4 *If $\overline{k} < k_0$, where k_0 denotes the coalition size associated with the intersection of the payoff function of the signatory and the non-signatory countries in the standard game, $\overline{k}$ is not a Nash equilibrium of the MPR-game.*

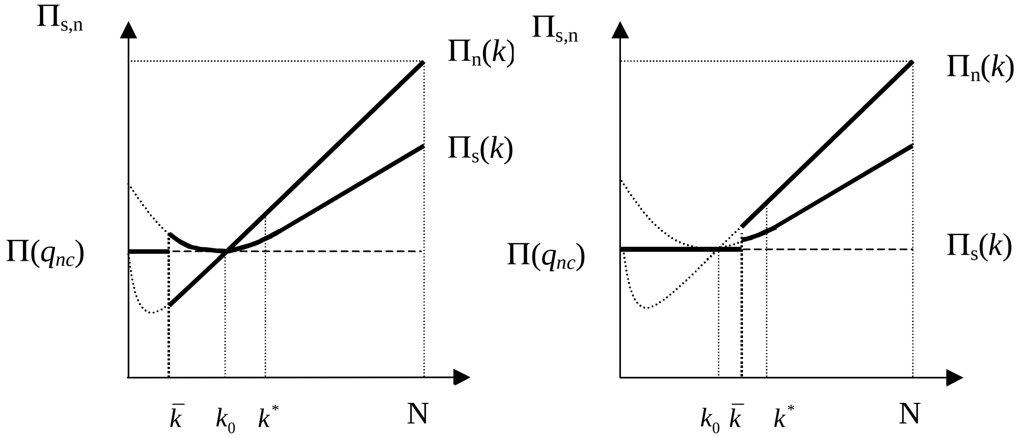

Figure 3: Non-binding minimumn participation levels

In the second case, when $\overline{k} \geq k_0$, $\overline{k}$ is a potential equilibrium. Potential, because this will depend on the size of the stable coalition (k^*) in the standard game. If $\overline{k} < k^*$, the minimum participation level will not be binding and the possibility that more countries than required ratify the agreement cannot be excluded.[9] This is intuitively clear, since, as depicted in the right picture of figure 3, the jump in the payoff functions occurs before k^* and therefore, the "old" equilibrium, implying higher payoffs to the signatory and the non-signatory countries than at $\overline{k}$, is still feasible for the negotiating countries. In other words, since k^*, the smallest possible stable coalition in the standard game, is still an equilibrium of the MPR-game, the stability condition (Eq. (1)) cannot be satisfied for $\overline{k} < k^*$. Note that in this case only modest efficiency gains are feasible, since a stable coalition of the size k^* implies that all countries choose an abatement level near the non-cooperative one, as discussed in section 3. These findings lead to proposition 5.

[9] Since $\overline{k}$ is by definition an integer number, this case is of course only possible if k^* is bigger then k_0 rounded up to the next integer number: $k^* > [k_0] + 1$. Rutz and Borek (2000) prove that such situations may arise.

Proposition 5 *If $k_0 \leq \overline{k} < k^*$, where k^* denotes the subgame perfect Nash equilibrium of the standard game, $\overline{k}$ is not a Nash equilibrium of the MPR-game. k^* is however still a subgame perfect Nash equilibrium of the MPR-game.*

Consider next the case $\overline{k} \geq k^*$, as depicted in figure 2. To see whether $\overline{k}$ constitutes an equilibrium of the MPR-game, we have to check if the two parts of the stability condition are satisfied. The left part of the stability condition, $\Pi_s(\overline{k}) \geq \Pi_n(\overline{k} - 1)$, is clearly satisfied: If a signatory country leaves the coalition, no agreement will be reached and all countries get the non-cooperative payoff since $\Pi_n(\overline{k} - 1) = \Pi_{nc}$. To check if the right part of the stability condition is satisfied as well at $\overline{k}$, a more subtle argument is needed. By definition of the equilibrium, the stable coalition in the standard game (k^*), satisfies $\Pi_n(k^*) \geq \Pi_s(k^* + 1)$. Further, if k^* is the unique equilibrium of the standard game, every $k > k^*$ must satisfy the right part of the stability condition as well.[10] The intuition for this is simply that for every $k > k^*$ the right part of the stability condition holds but not the left and vice versa for every $k < k^*$. Therefore, given the uniqueness of the equilibrium in the standard game, $\overline{k}$ necessarily must satisfy the right part of the stability condition as well, i.e. $\Pi_n(\overline{k}) \geq \Pi_s(\overline{k} + 1)$. Consequently, there cannot be an incentive for more than $\overline{k}$ countries to join the coalition in the MPR-game.

It is important to note that such a binding minimum participation level implies efficiency gains. This follows straight from proposition 1-3, which tells us that for any coalition size bigger than k_0, the payoff to the signatory and non-signatory countries is strictly larger than in a situation without cooperation. We summarize this result in proposition 6.

Proposition 6 *If $\overline{k} \geq k^*$, $\overline{k}$ constitutes a subgame perfect Nash equilibrium in pure strategies for the MPR-game, i.e. the game with a minimum participation rule.*

The next question to address is, if there are other equilibria in the MPR-game. By definition of the minimum participation rule there are always $\overline{k} - 1$ possible

[10] In Rutz and Borek (2000), we were not able to exclude the possibility of functional specifications resulting in multiple equilibria. Put simply, it is however safe to assume uniqueness of the equilibrium, as long as we are in a classical public good situation, i.e. a situation where freeriding gets more profiable the more countries are cooperating.

coalitions which prevent the achievement of an agreement. For every of these coalitions countries will choose the non-cooperative abatement level and consequently, the resulting payoff for all parties will be the non-cooperative one. Further, every coalition size smaller than $\overline{k} - 2$, satisfies the stability condition with equality, since no country can gain or lose by joining, respectively leaving a coalition. In contrast to the standard game, "no cooperation" is therefore another equilibrium of the MPR-game. Note however that this equilibrium is strictly dominated by the above described equilibrium where $\overline{k}$ countries form a coalition, respectively k^* countries when $\overline{k} \leq k^*$. This result is summarized in proposition 7.

Proposition 7 *For $\overline{k} > 2$, "no cooperation" constitutes a subgame perfect Nash equilibrium in pure strategies of the MPR-game.*

The introduction of a minimum participation rule into the standard game can thus substantially change the outcome of negotiations. If the minimum participation level is binding ($\overline{k} \geq k^*$), real efficiency gains are feasible in the MPR-game. This in contrast to the standard game, where the small resulting gains from cooperation associated with the stable coalition k^* are solely due to an "integer effect", i.e. coalition size being a discrete variable. The reason for the efficiency gains in the MPR-game is that a minimum participation rule changes the social dilemma structure of multilateral environmental negotiations to a game of chicken. In this class of games there are generally many non-cooperative equilibria, each of which may be optimal and none of which is dominant, and the task of the players is to coordinate their actions to select one.[11] Thus, if countries can efficiently solve the coordination problem, i.e. any $\overline{k}$ of the N countries ratify the agreement, gains from cooperation will result. In this sense, minimum participation rules can be interpreted as "hand-tying" mechanisms. That is, countries, aware of the freerider incentive inherent in the abatement of transboundary pollution problems, consciously agree to minimum participation rules to deter freeriding.

[11] Many authors (see e.g. Carraro and Siniscalco 1998) describe the structure of the standard model as well as a game of chicken. This, because the equilibrium is typically characterized by partial cooperation between countries. This seems however inaccurate, since in a game of chicken real efficency gains are feasible, while in the standard model no real gains from cooperation are possible (see Rutz and Borek 2000).

5 A Hypothesis

Having determined the equilibria of the MPR-game, it is of course tempting to study these theoretical predictions by analyzing existing multilateral environmental agreements. In practice, not all of the parameters of the above discussed model are however observable: The number of ratifications required by the minimum participation rule ($\overline{k}$) and the actual number of ratifications ($\widetilde{k}$) of an agreement are in general observable, while k^*, the stable coalition of the standard game, is unfortunately not observable. Thus, when considering an existing environmental agreement, three constellations of the two observable parameters can occur:

- $\overline{k} = \widetilde{k}$: the number of actual ratifications matches the number of countries required by the minimum participation rule.
- $\overline{k} > \widetilde{k}$: there are less ratifications than required by the minimum participation rule.
- $\overline{k} < \widetilde{k}$: there are more ratifications than required by the minimum participation rule.

In what follows, we will shortly discuss the meaning and relevance of these three cases in the framework of the above introduced model and then formulate a hypothesis, which will be the subject of section 6.

For the first two cases the interpretation is straightforward: A match between $\overline{k}$ and $\widetilde{k}$ simply suggests that countries have efficiently solved the coordination problem. In this case it is likely that gains from cooperation are realized, although, compared to a situation with full cooperation between countries, the result will only be second best. The second case ($\overline{k} > \widetilde{k}$) implies that an agreement has not entered into force yet. In other words, if there are less ratifications than required by the minimum participation rule, countries have not (or not yet) efficiently solved the coordination problem.

Concerning the third case ($\overline{k} < \widetilde{k}$), two interpretations can be thought of: either there is a coordination problem or the minimum participation rule is not binding. If more than the required number of countries ratify an agreement due to the coordination problem, one would expect that such off-equilibrium behavior is corrected over the course of time. That is, one should observe withdrawals from an agreement, a step hardly ever seen in reality.[12] Therefore it seems reasonable to exclude this possibility.

The second interpretation is that minimum participation rules are not always binding. According to propositions 4-6, a non-binding minimum participation rule implies that $\overline{k} \leq k^*$, i.e. the minimum participation rule is set lower than the equilibrium in the standard game. To address the question whether such a situation is likely to occur, we have to determine the restrictiveness of the requirement that $\overline{k} \geq k^*$. Since k^* is a non-observable parameter, the only possibility to tackle this question is to consult numerical simulations. Barrett (1994) conducted such simulations for the case of linear marginal abatement benefits and costs. Working with 100 players, his results reveal that for a wide range of cost and benefit parameter values the stable coalition does not exceed 3 countries, i.e. cooperation is limited to a very small fraction of the negotiating countries. Only for very small cost-benefit-ratios bigger stable coalitions do emerge. A low cost-benefit-ratio implies however that there is no marked public good problem inherent in the abatement of transboundary pollution, since potential freerider benefits are small. In the limit case, when the cost-benefit-ratio equals zero, there is no public good problem at all and the non-cooperative equilibrium coincides with the social optimum, implying the formation of a "grand coalition", i.e. a coalition formed by all countries involved in the negotiations. Further, illustrative simulations by Barrett (1994) with only 10 players suggest that these results are not sensitive to the number of countries involved in the negotiations.

These results are derived assuming identical countries, an assumption which is not very realistic when considering real-life agreements. The sparse simulations working with heterogeneous countries seem however to lead to similar findings. Assuming two types of countries - 3 big ones and 97 small ones - and again working with linear marginal abatement benefits and costs, Barrett (1998) derives stable coalitions for different cost-benefit-ratios.[13] As in the case of identical countries,

[12] Withdrawals from agreements do sometimes occur in practice. These are however mostly connected to the fact that an agreement is superseded by another treaty.

[13] In his paper, Barrett is addressing the problem of climat change due to CO_2-emissions. In this framework, the 3 big countries represent the United States, the former USSR and the European Union, which account together for about one-half of global CO_2-emissions, while the remaining countries are assumed to be small CO_2-emitters.

stable coalitions involving more than 3 countries do only result for very low cost-benefit-ratios.

Therefore, if a multilateral environmental agreement is addressing a trans-boundary pollution problem with clear characteristics of a public good, it seems in a first step reasonable to assume that any minimum participation rule requiring more than 3 countries is binding. Given that this assumption is correct, the following hypothesis can be formulated:

Hypothesis 1 *If the MPR-game captures reality adequately and the minimum participation rule is binding, not (substantially) more than* $\overline{k}$ *countries should ratify a multilateral environmental agreement.*

To study this hypothesis, we evaluated the accession data of the 122 multilateral environmental agreements studied in section 2. Results of this study are presented in the following section.

6 A Simple Test of the New Model

The test of hypothesis 1 we carry out in this section is of very simple nature: We compare for the 122 environmental agreements the actual number of countries that have ratified a treaty with the number of countries required by the minimum participation rule. The requirements of the minimum participation rules were elicited by studying the treaty texts, while ratification data is available from CIESIN (see footnote 4). Since this data was however often outdated, we used, where available, other sources to complete the data set. These are the *Internet Guide to International Fisheries Law*[14], the *Yearbook of International Co-operation on Environment and Development*[15] and various websites of convention secretariats.

Thus, our test is straightforward whenever a minimum participation level is explicitly stated in a treaty text and not coupled to other requirements, as is the case for 81 agreements (see table 1). In cases where the participation rule is however coupled to other requirements, it is sometimes difficult to specify an exact

[14] URL: http://www.oceanlaw.net/.
[15] URL: http://www.ngo.grida.no/ggynet/.

minimum participation level. Therefore, the results we present in the following are not concluding and do not claim completeness.

Not all of the 122 agreements were suited for our purpose: 2 do not require minimum ratification and for another 4 agreements, no accession data was available. Further, since we only consider participation rules requiring more than 3 countries as potentially binding, we had to exclude another 18 agreements from the evaluation. It should however be noted that most of these treaties are very small, addressing 5 or even less countries. Another 8 agreements had to be excluded from the analysis since they require unanimous ratification by the parties and membership is limited. Therefore, in these cases, minimum participation level and the number of ratifications necessarily must match. Finally, 3 agreements were superseded before reaching the required number of ratifications and therefore never entered into force.

This leaves us with 87 agreements to study hypothesis 1. In figure 4 we depicted the excess cooperation, defined as the number of actual ratifications minus the minimum participation rule, for these 87 agreements.

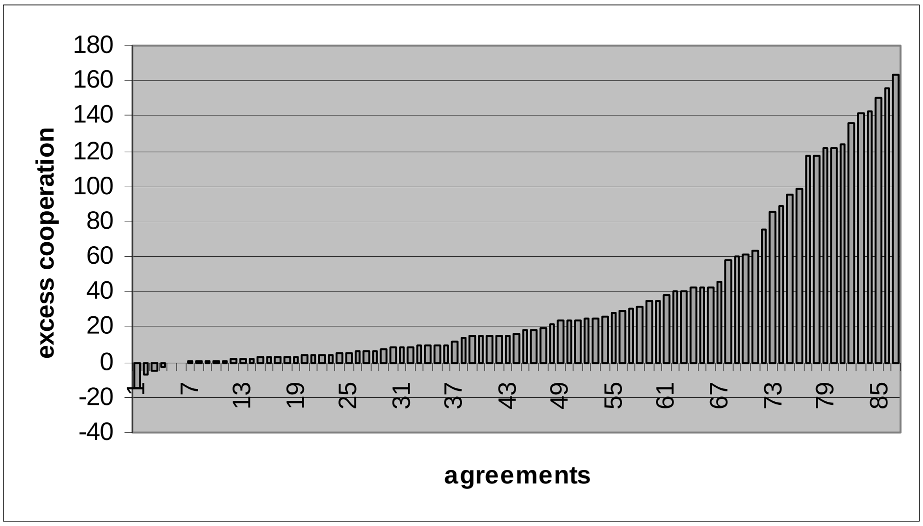

Figure 4: Excess cooperation in multilateral environmental agreements

The first thing to notice from figure 4 are the few cases of negative excess rates. These agreements, of course, have not entered into force yet. They illustrate however the fact that the above discussed scenario of less observed ratifications than required by the minimum participation rule ($\overline{k} > \widetilde{k}$) is not of much relevance in practice.

For the cases of positive excess rates, average excess cooperation amounts to approximately 40 countries, with a minimum of 0 and a maximum of 163 countries. "Excess cooperation" is however highly correlated to the (potential) size of an agreement. That is, such astonishing excess rates as 100 and more countries can only occur when considering truly global environmental problems, while for a small regional agreement even 3 ratifications in excess might prove substantial. It seems therefore somewhat arbitrary to set a threshold of the excess rate at which to consider hypothesis 1 as rejected. As figure 4 shows however, this poses no serious problem in many cases: For 40 agreements the excess rate amounts to more than 20 countries, which seems extremely high. In another 13 cases the excess rate still reaches between 10 and 19 countries. Accounting for the fact that none of these agreements are of global scope, all of them address approximately 20 to 50 countries, these excess rates still seem substantial.

When considering excess rates smaller than 10 countries, the threshold problem might however become important. If we set the threshold at 4 countries - as noted above, a rather arbitrary assumption - there are another 15 treaties contradicting our hypothesis. Again, it is important to note that none of these treaties are of global scope. There are thus 15 agreements left, for which the difference between the minimum participation level and the actual number of members accounts for less than 4 ratifications. A closer analysis of these agreements reveals however that in 7 cases this fact comes as no surprise: All these treaties are very small in scope, addressing between 6 and 8 countries, and the minimum participation level is set in a way that the excess rate cannot exceed 3 ratifications. In other words, 3 or less ratifications in excess are sufficient to integrate all potentially affected countries in the agreement.

The results of our study are summarized in table 2. Although these results, due to problems like e.g. coupled minimum participation rules and incomplete data, are not concluding, our study shows that even when setting a very restrictive threshold of 20 countries, about one-half of the considered 87 agreements do not confirm the theoretical predictions of the MPR-game.

excess cooperation	number of treaties
20 to 163 countries	40
10 to 19 countries	13
4 to 9 countries	15
0 to 3 countries	15
-15 to -1 countries	4
total	**87**

Table 2

Should we conclude from theses findings that the discussed model does not capture reality adequately? From a game-theoretic point of view, it is indeed not easy to answer the question why additional countries join a multilateral environmental agreement once it entered into force. A discussion of this puzzle and two suggestions for an extension of the MPR-game follow in the next section.

7 Explaining Accession Behavior

7.1 A Game-Theoretic Puzzle

Multilateral environmental negotiations in reality are of course no "one-shot" games as postulated by the discussed model. Rather, once an agreement entered into force, there is still the possibility for new accessions and such accessions are a widely observed phenomenon. A further observation seems interesting: After an agreement entered into force, every new acceding country will have to accept the existing terms of the treaty. In most agreements reservations are not possible and substantial amendments to treaties are rarely seen.

From a game-theoretic perspective, this would suggest that the discussed game should be seen as a dynamic game. To be more precise, such a dynamic game implies that in every period after the agreement entered into force, the non-signatory countries can decide if they want to join the agreement, given the terms of the treaty negotiated in the first period. However, if there is no possibility to change the terms of a treaty, there seems to be no reason why a country should accede an agreement once it entered into force, i.e. when already $\overline{k}$ countries ratified the treaty. Given the public good character of transboundary pollution, such a decision can never satisfy the criteria of individual rationality. Therefore, from a game-theoretic point of view, the observed accession behavior can only be explained if the freerider problem is not prevailing. Two explanations why under certain circumstances this freerider incentive might not be prevailing will be offered in the following. Both of them should be seen as potential extensions of the MPR-game, which does not explicitly model the negotiation process between countries. This process logically takes place before countries decide whether or not to ratify a treaty text and might crucially influence the feasible efficiency gains of an agreement. In other words, the above discussed game implicitly assumes that a treaty text is drafted in a way to ensure the achievement of a first best solution in the case of full cooperation between countries. Strategic behavior in the negotiation process can however result in a situation where even in the case of full cooperation only second best solutions are feasible.

7.2 Diluted Terms of a Treaty

The first explanation for the observed accession behavior would be that through the above mentioned negotiation process, the terms of a treaty get so diluted that the imposed obligations are not much stricter than what a country would have done anyway when pursuing an unilateral policy. This, because no country can be forced to accede an agreement and therefore, any potentially successful treaty text must be drafted in a way to find the approval of at least $\overline{k}$ of the negotiating parties. Taking into consideration the specific demands and requests of all involved countries however often leads to the adoption of a policy of the "least common denominator", which can result in a severe dilution of the terms of a treaty.

As noted in section 5, in such a situation, the freerider incentive is relatively weak, since diluted terms of a treaty imply a low cost-benefit-ratio. Then, even in the standard game big stable coalitions can emerge and it is therefore likely that a minimum participation rule no longer is binding. Thus, once an agreement entered into force, i.e. the required minimum ratifications are reached, a non-signatory country does not incur substantial costs when additionally joining the agreement. For a government of a non-signatory country such a decision might even be profitable by building up a reputation as a "green government".

This scenario would of course imply that multilateral environmental agreements with more signatory countries than required by the minimum participation rule tend to be ineffective. That is, they do not result in a substantial improvement of environmental quality compared to a situation without cooperation between countries.

7.3 Endogenous Minimum Participation Rules

The second explanation is no less pessimistic. As noted in section 4.1, minimum participation rules are in practice of course not determined exogenously, but an integral part of the negotiation process. To account for this, the minimum participation level would have to be endogenized in the above game. Two effects can be expected from such an extension of the game: On the one hand, by setting a high minimum participation level, countries can minimize the freerider incentive, which increases potential gains from cooperation. On the other hand, however, since the discussed game is basically a coordination game, a high minimum participation level increases the possibility of a coordination failure, which implies that the agreement will not enter into force and all countries get the non-cooperative payoff. In addition and probably even more important than the coordination problem, is the fact that high minimum participation levels increase the bargaining power of the individual country. This, because threats to jeopardize an agreement get more credible. In the limit case when a minimum participation rule requires unanimity, it lies in the power of every single negotiating country to prevent the achievement of an agreement. Thus, there seems to be a trade-off between setting too high and too low minimum participation levels.

If this trade-off results in countries choosing a minimum participation level $\overline{k}$ near or equal to k^*, the stable coalition in the standard game, the observed accession pattern can be explained: In the first period $\overline{k}$ countries ratify the agreement. But since $\overline{k} \approx k^*$, all signatory and non-signatory countries commit to an abatement level near the non-cooperative one, as discussed in section 3.2. Again, as with the first explanation, countries joining the agreement after its entry into force will consequently have to accept obligations which are not much stricter than what a country would have done unilateral. The implications of this scenario are the same as with the first: The reason why additional countries join multilateral environmental agreements after the entry into force lies in the tendency to ineffectiveness of such treaties.

However, the effects of strategic behavior of country representatives in the negotiation process on the outcome of negotiations is not well investigated yet. The extensions of the discussed model, as presented in this section, are not concluding, but should be seen as suggestions for further research on the subject.

8 Summary and Conclusions

The goal of this paper was to assess the effects of minimum participation rules on the outcome of multilateral environmental negotiations. Although such rules are widely used in practice, up to now, minimum participation levels have hardly been taken into consideration when modelling negotiations on transboundary pollution problems. By incorporating a minimum participation rule in a widely used game-theoretic model of coalition formation, we have shown that this can however crucially influence the outcome of negotiations. When a minimum participation rule is binding, i.e. the required number of ratifications is set higher than the number of countries that would have acceded an agreement anyway, real efficiency gains are feasible. This, in contrast to the standard game, where no efficient cooperation between countries can be expected.

Besides the theoretical implications, the MPR-game allows for a simple prediction about how many countries will ratify multilateral environmental agreements in practice: Given that the discussed model captures reality adequately, one would expect that not more than the number of countries required by the minimum participation rule ratify an agreement. Our study reveals however that in most cases substantially more then predicted the number of countries ratify international environmental agreements.

From a game-theoretic point of view, the question why additional countries join a multilateral environmental agreement once it entered into force, leaves us with a puzzle. We offered two explanations for this phenomenon. Both of them are possible extensions of the discussed model and seem to suggest that many existing multilateral environmental agreements are not very successful, i.e. they do not result in a substantial improvement of environmental quality compared to a situation without an agreement. Given that the chosen game-theoretic approach reflects reality adequately, the observed accession pattern thus suggests a high degree of ineffectiveness of many multilateral environmental agreements.

This suggestion is in line with the sparse empirical economic literature on the effectiveness of multilateral environmental agreements. As far as the authors are aware, there exist only two empirical studies addressing this subject. The first is by Murdoch and Sandler (1997), analyzing the Helsinki Protocol, while the second, by Finus and Tjotta (1998), addresses its successor the Oslo Protocol. Both studies conclude, that the abatement targets laid down in the protocol do not go beyond what countries would have done in their self-interest anyway. Although the simple test we conducted in this paper cannot be considered as empirical evidence, our results seem to support the findings of these empirical studies. Further research on this subject will however be needed.

References

[1] S. Barrett, International Environmental Agreements as Games, *in* "Conflicts and Cooperation in Managing Environmental Resources" (R. Pething, Ed.), Springer-Verlag, Berlin (1992).

[2] S. Barrett, Self-enforcing international environmental agreements, *Oxford Econom. Papers* **46**, 878-894 (1994).

[3] S. Barrett, Towards a theory of international environmental cooperation, *in*: "New directions in the economic theory of the environment" (C. Carraro and D. Siniscalco, Ed.), Cambridge University Press (1997).

[4] S. Barrett, Heterogeneous international environmental agreements, *in*: "International Environmental Agreements: Strategic Policy Issues" (C. Carraro, Ed.), E. Elgar, Cheltenham (1998).

[5] J. Black, M.D. Levi and D. de Meza, Creating a Good Atmosphere: Minimum Participation for Tackling the "Greenhouse Effect", *Economica* **60**, 281-293 (1993).

[6] M. Botteon and C. Carraro, Burden sharing and coalition stability in environmental negotiations with asymmetric countries, *in*: "International Environmental Agreements: Strategic Policy Issues" (C. Carraro, Ed.), E. Elgar, Cheltenham (1998).

[7] C. Carraro and D. Siniscalco, Strategies for the international protection of the environment, *Journal of Public Economics* **52**, 309-328 (1993).

[8] C. Carraro and D. Siniscalco, International environmental agreements: Incentives and political economy, *European Econom. Rev.* **42**, 561-572 (1998).

[9] A. Endres, M. Finus und F. Lobigs, Symbolische Umweltpolitik im Zeitalter der Globalisierung? - Zur Effektivität internationaler Umweltverträge aus ökonomischer Sicht, *Perspektiven der Wirtschaftspolitik* **1**, 73-91 (2000).

[10] M. Finus and S. Tjotta, The Oslo Agreement on Sulfur Reduction in Europe: The Great Leap Forward?, *Working Paper*, University of Bergen, Bergen (1998).

[11] M. Hoel, Efficient Climate Policy in the Presence of Free Riders, *J. Env. Econom. Management* **27**, 259-274 (1994).

[12] J.C. Murdoch and T. Sandler, Voluntary Cutbacks and Pretreaty Behavior: The Helsinki Protocol and Sulfur Emissions, *Public Finance Review* **25**, 139-162 (1997).

[13] S. Rutz und T. Borek, International Environmental Negotiations: Does Coalition Size Matter?, *Working Paper Series,* Center for Economic Research, Swiss Federal Institute of Technology, **20/00**, Zürich (2000).