CICERO Working Paper 2004:06

# The Theory of Full International Cooperation

## An Experimental Evaluation

**Leif Helland and Jon Hovi**

June 2004

**CICERO**
Center for International Climate
and Environmental Research
P.O. Box 1129 Blindern
N-0318 Oslo, Norway
Phone: +47 22 85 87 50
Fax: +47 22 85 87 51
E-mail: admin@cicero.uio.no
Web: www.cicero.uio.no

**CICERO Senter for klimaforskning**
P.B. 1129 Blindern, 0318 Oslo
Telefon: 22 85 87 50
Faks: 22 85 87 51
E-post: admin@cicero.uio.no
Nett: www.cicero.uio.no

**Tittel: The Theory of Full International Cooperation: An Experimental Evaluation**

**Forfatter(e):** Leif Helland og Jon Hovi
CICERO Working Paper 2004:06
22 sider
**Finansieringskilde:** Norges forskningsråd
**Prosjekt:** Håndheving Kyoto (0408); Håndheving, verifikasjon, og design av klimaavtaler (1410)

**Prosjektleder:** Jon Hovi
**Kvalitetsansvarlig:** Gunnar Eskeland
**Nøkkelord:** Spillteori, reforhandlingssikker likevekt, eksperimenter, Kyotoprotokoll, etterlevelse, håndheving

**Sammendrag:**. Ideen om reforhandlingssikker likevekt har etter hvert blitt noe av en hjørnestein i nyere spillteoretisk basert forskning om internasjonale miljøavtalers stabilitet. Ved å anvende dette løsningsbegrepet på en lineær versjon av uendelig gjentatte Fangens Dilemma-spill har Scott Barrett utledet et antall interessante (og til dels provoserende) prediksjoner om internasjonalt samarbeid for å begrense globale klimaendringer. Denne artikkelen rapporterer resultatene fra et laboratorie-eksperiment som ble gjennomført for å teste to sentrale prediksjoner fra Barretts arbeider. Den første prediksjonen sier at desto høyere kostnad som påløper ved å samarbeide, jo flere land vil gjøre dette. Den andre sier at antallet samarbeidende parter er uavhengig av gruppestørrelse. Eksperimentet ble designet for å tilnærme forutsetningene i Barretts modell så godt som mulig. Konklusjonen er likevel at resultatene gir svært begrenset støtte til de to nevnte prediksjonene.

**Språk:** Engelsk

**Title: The Theory of Full International Cooperation: An Experimental Evaluation**

**Author(s):** Leif Helland and Jon Hovi
CICERO Working Paper 2004:06
22 pages
**Financed by:** The Research Council of Norway
**Project:** Enforcing the Kyoto Protocol (0408); Compliance, verification, and design of climate treaties
**Project manager:** Jon Hovi
**Quality manager:** Gunnar Eskeland
**Keywords:** Game theory, renegotiation-proof equilibrium, experiments, Kyoto Protocol, compliance, enforcement

**Abstract:** The concept of renegotiation-proof equilibrium has become a cornerstone in recent game theoretic reasoning about the stability of international environmental agreements. Applying this solution concept to a linear version of the infinitely repeated N-person Prisoners' Dilemma, Scott Barrett has been able to derive a number of interesting (and sometimes provocative) predictions about international cooperation to curb climate change. This paper reports the results of a laboratory experiment designed to test two central predictions from Barrett's model. The first prediction says that the higher the cost of making a contribution, the more cooperation will materialize. The second claims that the number of cooperators is independent of group size. The experiment was designed to replicate the assumptions of Barrett's model closely. We find that the experimental confrontation lends very little support to the two predictions.

**Language of report:** English

Rapporten kan bestilles fra:
CICERO Senter for klimaforskning
P.B. 1129 Blindern
0318 Oslo

Eller lastes ned fra:
http://www.cicero.uio.no

The report may be ordered from:
CICERO (Center for International Climate and Environmental Research – Oslo)
PO Box 1129 Blindern
0318 Oslo, NORWAY

Or be downloaded from:
http://www.cicero.uio.no

## Contents



## Acknowledgements

We gratefully acknowledge helpful comments by Geir Asheim, Scott Barrett, Ernst Fehr, Urs Fischbacher, Michael Kosfeld and Lorenz Götte. Urs Fischbacher also produced practical help and advice in programming the experiment. Lynn A. P. Nygaard provided excellent editorial assistance. Colleagues at CICERO kindly participated in a pilot project, and provided helpful comments afterwards. The project of which this paper is a part received financial support from the Research Council of Norway.

# 1 Introduction

The concept of a renegotiation-proof equilibrium has become a cornerstone in game theoretic reasoning about the stability of international environmental agreements in general and climate agreements in particular.[1] This concept rests on a particular notion of collective rationality. Loosely put, the idea is that a punishment that hurts *all* players will not be carried out. Unless an agreement is renegotiation proof, a transgressor will be able to escape a threatened punishment by inviting the other players to renegotiate, rightfully pointing out that every player gains by skipping the punishment. Foreseeing this, rational decision makers will not enter into agreements that are susceptible to renegotiation.

This notion of collective rationality was introduced into the study of environmental politics by Scott Barrett (e.g., Barrett 1994, Barrett 1999a, Barrett 2002, Barrett 2003). Barrett's work has attracted considerable attention. One reason for this is that it is (rightly) seen as a significant addition to the literature on public goods provision and international cooperation. A second reason is that Barrett's work offers some rather distressing predictions. In particular, Barrett predicts that within a given group of states, only agreements of *little* value to the group are compatible with participation by all members of the group. By contrast, agreements of *considerable* value to the group will be implemented by only a small number of the states. This represents a rather pessimistic view of the prospects for effective and global cooperation to curb climate change.

While Barrett applies a variety of theoretical models, some of his most provocative predictions have been deduced from a particular version of the infinitely repeated N-person Prisoners' Dilemma. Barrett is careful to point out that this model is based on a number of rather restrictive assumptions.[2] Nevertheless, there is no doubt that the model is taken seriously in the literature, at least as a benchmark. A number of scholars have used this benchmark as a starting point for discussing in detail additional measures to secure compliance with international agreements (Barrett 2001; Barrett 2003; Finus 2001; Finus & Rundshagen 1998a b; Carraro 1999; Carraro & Siniscalco 1998; Wagner 2001). Among the measures typically discussed in this literature are taxes, quotas, regional agreements, trade sanctions, technology standards, transfers and other issue linkages.

In this paper, we focus on the benchmark model itself. More specifically, we report the results of an experiment carried out in a laboratory setting. Obviously, no laboratory experiment can even come close to mimicking all the intricacies of international politics. Nor is this something that the benchmark model itself is able (or even attempts) to do. However, the experiment was carefully designed to replicate the environment of the model. Interestingly, the experimental confrontation lends little, if any, support to the central predictions of the benchmark model. We conclude that our results challenge the empirical relevance of this model, which underlies a good deal of recent theorizing on international climate agreements.

---

[1] For example, see Barrett (2002, 2001, 1999a,b, 1994); Carraro (1999); Carraro & Siniscalco (1998); Ecchia & Mariotti (1998); Enders & Finus (2002); Finus (2001); Finus & Rundshagen (1998a b); Wagner (2001).

[2] Among the assumptions are the following: (i) All countries are identical; (ii) interaction takes place under complete and (almost) perfect information; (iii) cooperation exhibits constant or increasing returns; (iv) a uniform emission quota is the only instrument available; (v) abatement levels from previous periods are observed noiselessly at the beginning of every period, at no cost; (vi) punishments can be carried out with full force immediately after observing abatement levels; (vii) the only possible sanction to non-compliance is to nullify one's own abatement efforts; (viii) cost functions are independent (e.g., interaction effects via world markets are not taken into account); (ix) the choice of abatement levels is binary (either socially efficient or no abatement).

## 2 The model

In Barrett's model,[3] and in most variants of it, curbing global warming is seen as a pure public good. Global warming is caused by emissions of greenhouse gases. Such gases mix very quickly and almost perfectly in the atmosphere. Assuming that the beneficial consequences of reduced warming are spread evenly, a costly reduction in emissions will benefit any country regardless of its geographical location.[4] The upshot is that every country faces a free-rider incentive.

As already mentioned, the model is a special case of the infinitely repeated N-person Prisoners' Dilemma (PD), which is well known from numerous previous theoretical and experimental studies. In particular, this game is frequently used to study the conditions for provision of pure public goods. However, two characteristics of Barrett's model deserve particular attention. First, he assumes that the utility curves are linear functions of the number of cooperating parties. Second, he uses the notion of renegotiation-proof equilibrium as solution concept.[5] More specifically, he analyses a situation where all parties use a strategy called "Getting Even." This strategy instructs party $j$ to cooperate in a given round unless $j$ has defected less often than any other party in previous rounds. Given that cooperation is based on Getting Even, a defection will trigger all the other players to defect in the next round. Cooperation will be reestablished once the defector has "accepted" the punishment by being the only party to cooperate in one round of the game.

We find it useful to introduce some notation. Assume that there are N identical players, and that in each round of the game each player must choose between Contribute (reduce emissions of greenhouse gases) and Abstain (not reduce emissions). Let the periodic utility to a player of contributing to the collective good be $U(\text{Contribute}) = d \cdot m - c$, and let the corresponding utility to the player of not contributing be $U(\text{Abstain}) = b \cdot m$. In these expressions $d$ and $b$ stand for the marginal gain of contributing or not contributing respectively, $c$ is the cost of a contribution, and $m$ is the number of contributors (smaller than or equal to N). For the stage game to be a Prisoners' Dilemma, we must have that $U$(Contribute) is smaller than $U$(Abstain) regardless of the number of other cooperators. Furthermore, $U$(Contribute) for $m = \text{N}$ must be greater than $U$(Abstain) for $m = 0$.

Every renegotiation-proof equilibrium is a subgame perfect equilibrium, and every subgame perfect equilibrium is a Nash equilibrium. The precise conditions under which Getting Even forms a Nash equilibrium and a subgame perfect equilibrium are set out elsewhere (Barrett 1999, Finus & Rundshagen 1998a, Farell & Maskin 1989). We concentrate on the conditions under which a situation where all players are using the Getting Even strategy forms a renegotiation-proof equilibrium. For clarity of exposition we limit the

---

[3] For convenience and linguistic variation, we use "the benchmark model" and "Barrett's model" interchangeably. We do not imply that all (or even most) of Barrett's work is based on the particular model under scrutiny in this paper. Rather, the present model is one of several models studied by Barrett.

[4]The assumption of evenly spread beneficial consequences might seem strong. However, little is known about the distributions of benefits (costs) following a reduction (an increase) in the average global temperature (UNEP2001). Furthermore, not much has been learned about this over the last 10 years (Schelling 1992:2-3 compare UNEP2001). This justifies the assumption on the principle of insufficient knowledge.

[5]Several competing concepts of renegotiation-proof equilibrium exist in the literature (see Bergin and MacLeod 1993 for a discussion). When we use the concept in this article we always refer to the concept of weakly renegotiation-proof equilibrium, formalized for 2×2 games by Farrell and Maskin (1989), and extended to N-person games by Barrett (1994) and by Finus and Rundshagen (1998). The concept is discussed in greater detail in a subsequent section of this paper.

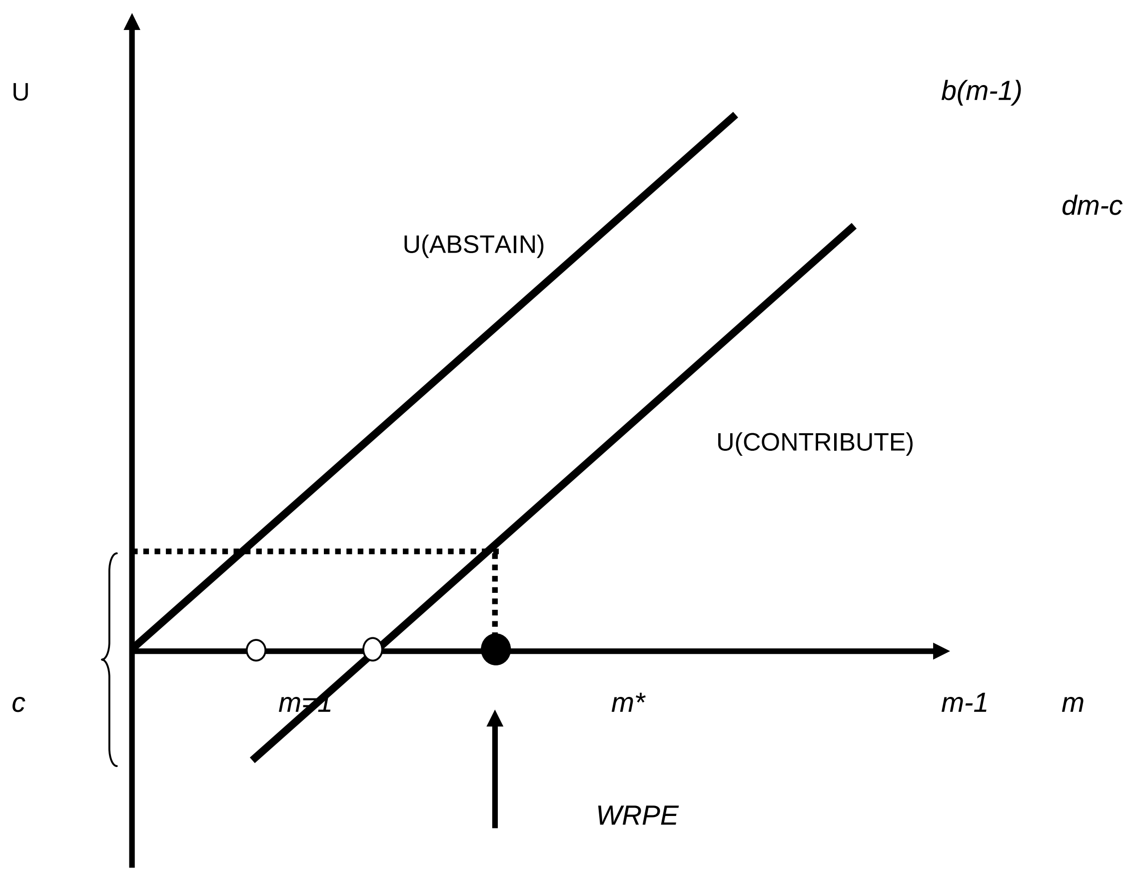
discussion to the case where the players are perfectly patient (discount factors arbitrarily close to one).

Assume that m* players have agreed to contribute. We call these m* players "signatories." Suppose one of the signatories deviates from the agreement in period $t$ and goes back to Getting Even from period $t + 1$ onwards. According to the dictums of Getting Even, the deviating party must pay penance by being the only party to cooperate in period $t + 1$. This nets each of the other signatories $b \cdot 1$ in that period. By contrast, if these signatories collectively decide to ignore the deviation (in contradiction with the dictums of Getting Even), their payoffs will be $d \cdot (m^*) - c$ in period $t + 1$. Regardless of whether other signatories decide to punish the deviating party or to renegotiate, cooperation will be restored from period $t + 2$ onwards. Thus, it is individually rational to carry out the punishment if $b \geq d \cdot (m^*) - c$. In other words, the agreement is renegotiation proof only if $m^* \leq (b + c)/d$.

We now see that the requirement of renegotiation-proofness imposes a restriction on the maximum number of signatories to an agreement. Note also that this maximum is independent of the player set (N). Furthermore, a less valuable agreement can sustain more signatories than one that is more valuable. For instance, other things being equal increasing the cost of making a contribution ($c$) increases the maximum number of signatories to the agreement.

Thus, with linear payoff functions it follows that there must be an upper limit to the number of players that cooperate in equilibrium. This maximum is small if the gap between the payoff curves is narrow. It becomes larger if the gap is widened, i.e., if the cost of cooperation increases. This aspect of the model comes clearly across in the algebra above, and is illustrated graphically in figure 1.

**Figure 1: The Stage Game**



A number of very strong predictions can be derived from the model. In general, model predictions fall into two categories: point predictions and directional predictions. A point prediction indicates the exact value of a certain (dependent) variable, given particular values of a set of other (independent) variables. A directional prediction is less ambitious, and merely suggests the direction of change in a dependent variable likely to be caused by a particular change in some independent variable (e.g., Hovi, Sprinz and Underdal 2003).

As Fiorina (1996) reminds us, only if all relevant (economic, political, institutional, demographic, technological, normative, cultural, and psychological) factors were fully incorporated into the model, could we hope to make a successful point prediction. However, "since we can never know all of the relevant variables, let alone measure all of them accurately, the empirical predictions of all models in political science – not just R[ational] C[hoice] – are about relative differences and comparative statics" (Fiorina 1996:88). This means is that we cannot discard a model simply because some (or even all) of its point predictions fail to get empirical support. A more reasonable test is to check whether the model's (main) *directional* predictions get empirical support. Following this methodological principle, we focus on two main predictions that can be derived from Barrett's model:

*PROPOSITION 1: The higher the cost of making a contribution, the more cooperation will materialize.*

Most readers are likely to find this hypothesis strongly counterintuitive. Surely, not many would expect that increasing the cost of a contribution makes cooperation *more* attractive? Consider the debate over which instruments efficiently curb climate change. Economists have emphasized the efficiency of markets for tradable emission quotas. Such markets help reduce abatement costs, and decision makers have taken the advice seriously. As a consequence, a number of countries will shortly be operating on such markets. It follows from the benchmark model that such markets *reduce* the number of countries willing to abate.

*PROPOSITION 2: The number of cooperators is independent of group size.*

Within the literature on public goods provision, several authors have focused on the impact of group size on cooperation (e.g., Olson 1965, Chamberlin 1982). By contrast, in Barrett's model the number of contributors in equilibrium depends only on the slopes of the utility curves and the distance between them (the cost of cooperation). Therefore, the number of parties that contribute in equilibrium does not depend on the overall number of players. Thus, according to the model, increasing group size has no impact on the amount of cooperation.[6]

Testing propositions one and two in a carefully controlled experiment represents a check of the Barrett's model. A fairly strong case for experimental evaluation of refinements can be made. As Larry Samuelson makes clear, the series of attempts at evaluating various refinements from within game theory has been no success story (1997:9): "The result has

[6]With an alternative concept of collective rationality Barrett (2002) establishes a link between equilibrium behaviour and the size of the player set. This alternative concept is less clear than the renegotiation-proof equilibrium considered here, and will not be discussed further.

been a race between those who construct examples to highlight the flaws of existing refinements and those who construct refinements. This race has produced an ever-growing collection of contending refinements, many of which are quite sensitive to fine details in the construction of the model, but the race has produced very little basis for interpreting these refinements or choosing between them."

Using experimental data to check the sensibility of refinements is not a novel idea.[7] Nonetheless, experimental evaluation of renegotiation proofness in the open horizon N-person PD-game has to our knowledge not been carried out. In this sense the results contained in our paper represent something new.

## 3 Previous research

The idea of a renegotiation-proof equilibrium for repeated games was formulated by Farrell (1983) and formalized by Farell and Maskin (1989) for 2 × 2 games. In particular, Farrell and Maskin identify the restrictions on discounted payoffs imposed by the requirement that the strategy vector be renegotiation proof (1989:335ff.). Van Damme (1989) considers the strategy vector where both players play Getting Even in a symmetric Prisoners' Dilemma game with an open horizon. He demonstrates that this vector constitutes a renegotiation-proof equilibrium for sufficiently high discount factors. More generally, he shows that the requirement of renegotiation-proofness does not further restrict the set of equilibrium payoffs, compared to the restriction of subgame perfection.[8] Barrett (1994) verifies that the Getting Even strategy vector also forms a renegotiation-proof equilibrium in an N-person Prisoners' Dilemma game with an open horizon, provided that the discount factors are arbitrarily close to unity. Moreover, in the N-person game, renegotiation-proofness does have more restrictive power on the set of equilibrium payoffs than subgame perfection. In particular, as we have already seen, renegotiation-proofness places a cap on the maximum number of parties to an agreement. Finally, Finus & Rundshagen (1998a) extends the N-person argument to the case where discount factors are *not* arbitrarily close to unity. As argued above, the restrictions imposed by renegotiation proofness in the N-person PD game are rather counterintuitive and invite empirical evaluation.

A massive amount of experimental research has been carried out on public goods games – of which the Prisoners' Dilemma is one variant. Almost all the experiments implement a fixed horizon of either one or several rounds that is known to the subjects *ex ante*. Reviews of the literature are given by Dawes and Thaler (1988), Ledyard (1995), and by Fehr and Schmidt (1999:836-9). Main findings from these experiments can be summarised as follows. Public good games with a fixed horizon have a unique equilibrium, in which no player contributes to the public good. In single-round experiments, average contributions lie in the interval 40–60 percent. When the public goods game is repeated – usually for 10 or 20 rounds – average contributions in general start out as in the one-shot game, but drop with time (though exceptions have been documented). The last period of the repeated game in general exhibits low levels of average contributions (high levels of free riding). When restarted, or played by subjects with previous experience from public goods games, however, average contributions start out as before, and in general evolve as described. Hence, the observed pattern cannot easily be attributed to learning. Lumpy goods (provision thresholds) seem to increase the average level of contributions, but this effect is uncertain. Pre-play communication tends to

[7]For instance, carefully controlled experiments aimed at evaluating various refinements in signalling games has been carried out. Se Banks *et al.* (1994), Brandts & Holt (1992, 1993, 1995) and Partow & Schotter (1996), Potters & van Winden (1996, 2000). Confer also Camerer (2003) for a thorough discussion.

[8]Thereby proving wrong the conjecture put forward by Farrell (1983) that only repetition of the stage game equilibrium of a discounted PD game under an open horizon could be renegotiation proof.

increase the average contribution levels. Increasing the marginal per capita return on a contribution has a strong positive effect on average contribution levels.[9] Controlling for this effect, the effect of group size is somewhat uncertain, but, if anything, average contribution levels tend to increase with increasing group size.[10] High costs (low benefits) of contributing tend to decrease the average contribution level, other things being equal.[11]

As far as we know, the only public goods experiment with an open horizon (i.e., where the players cannot identify a final period with certainty) is described in Roth and Murnighan (1978). This is remarkable, since the bulk of theoretical research has focused on open horizon games, where equilibrium contribution levels generally are higher than in fixed horizon games. The Roth and Murnigham study does not enable us to decide whether contributions tend to fall over time. Their main finding is that reducing the continuation probability reduces average contributions in the game.

The aim of Roth and Murnighan is different from ours. They explore whether variation in continuation probabilities (discount factors) has an impact on behaviour. Costs and benefits of the stage game are kept constant over treatments, and the study is limited to a two-player game (in which one of the players is an automaton). Furthermore, the generality of their results may be questioned because of the reward structure implemented. Subjects were rewarded solely by an opportunity to win 10 USD in a final game. In our experiment the continuation probability  is kept constant, while the cost of making a contribution is systematically varied over treatments. We study an N-person game (with no automatons), and rewards are accumulated from combinations of actions taken over the rounds of play.

We have not come across any experiment designed to test the notion of renegotiation-proof equilibria in repeated PD games, or – more generally – in public goods games. Given the apparent enthusiasm for this equilibrium concept in the literature, it seems worthwhile to fill this gap. In light of the fact that the concept is widely employed in research on a future global climate agreement, experimental confrontations seem particularly appropriate. This is so since a global climate agreement has not (yet) been implemented. Testing the implications of theoretical models on field data is therefore not (yet) a viable option.[12]

---

[9]The marginal per capita return must not be confused with the marginal gains of contributing or abstaining (the parameters *b* and *d* in Barrett's model). Write the subject's utility function on the following form:

$U_i = f(z_i-c_i)+g/N(c_i + \sum_{j\neq i} c_j)$, where $z_i$ is the private wealth of subject *i*, $c_i$ is the contribution of *i*, and N is the number of players. The optimal contribution to the good is given by $U_i'(c_i)=0 \Rightarrow f=g/N$, where *f* is the marginal return of resources used on the private good, while *g*/N is the marginal per capita return on resources used on the collective good. In our experiment $z_i\, \varepsilon \{6, 12, 21\}$ depending on the treatment, $c_i\, \varepsilon \{0, z_i\}$, $f = 1$ and $g = 3$. So the marginal per capita return is 3/5 if N=5 and 3/10 if N=10. With $f = 1$ it is a dominant strategy in the stage game not to contribute if $g/N<1$. This is the case in our experiment. As we can see, it is not possible to keep *d* and *b* constant (as our design demands), and at the same time hold the marginal per capita returns constant while increasing N.

[10]A question currently investigated on the research front is to what extent different kinds of "social preferences" can account for anomalies observed in the lab. Overviews of this research can be found in Ledyard (1995) and in Gächter & Fehr (1999).

[11]See especially Isaac, McCue & Plott (1985) and Marwell & Ames (1979).

[12]As far as we know, the only example of empirical research on the level of cooperation with a global climate agreement is Fredriksson & Gaston 2000. They estimate the conditional probability of countries ratifying the UN Framework Convention on Climate Change. The study does not address the content of the convention, or the chances that signatories to the convention will eventually comply with it.

## 4 Experimental design

*Recruiting subjects*: The experiment was undertaken over two consecutive days, with 20 students participating each day. To recruit subjects, posters were placed at a number of locations throughout the campus of the University of Oslo. Altogether, we received e-mails from more than 50 students who signalled an interest in participating. Around half were political science students. The rest came from various other departments. From the set of volunteers, 10 political science students and 10 students from other departments were randomly drawn to participate in the experiment each day. In the event that some participants would not show up, we also invited three potential substitutes for each day. Only one of these was actually used.

*Questionnaire 1:* Upon arrival, the students were asked to fill out a questionnaire asking them about their name, sex, age, university courses, and prior knowledge of game theory. They were also asked a few questions intended to tap potentially relevant aspects of their value system.

*Instructions and trial runs:* Together with the questionnaire, the subjects received written instructions explaining the details of the experiment. They also received a set of payoff matrixes. Included in the instructions was a set of control questions that enabled us to verify that the students had properly understood the payoff matrixes.[13] Before starting the experiment, two trial runs were executed to enable the subjects to get familiar with the software. The trial runs did not involve monetary payoffs.

*Experimental runs:* The experiment was conducted in a number of sessions. In each session, each of 2 or 4 groups played an infinitely repeated N-person Prisoners' Dilemma with linear payoff functions. Each group consisted of 5 or 10 subjects. After each round of play, a random device included in the software decided with probability 0.9 that the game would proceed to another round, and with probability 0.1 that the game would stop. The resulting number of rounds varied from 1 to 48, for an average of 12.3 rounds.

Throughout the experiments, we systematically varied group size, the cost of cooperation, and the information communicated to the students. On day 1 the game was played in four groups of five subjects. In sessions 1–3 the cost of cooperation was "low," whereas in sessions 4–5, the cost was "high." After the fifth session Barrett's proposed solution was explained in some detail to the subjects. The subjects were also allowed to ask questions about the nature of the solution. Then two new low-cost and two new high-cost sessions were undertaken. On day 2 this basic design was repeated (with a new set of subjects). However, there were two important differences. First, the subjects now played in two groups of 10 (as opposed to four groups of 5 on the first day). Second, additional sessions with a "very high" cost were included both before and after the students were informed of Barrett's proposed solution.

*Anonymity over sessions:* To make sure subjects were unable to condition their behaviour on the history of previous sessions, we used a double randomization. First, subjects were randomly distributed to groups before each new session. Second, subjects were randomly assigned a participant number in the group at the beginning of each new session (a number between 1 and 5 in the small group treatment, and a number between 1 and 10 in the large group treatment).

*Context*: It is well known that the context imposed on an experiment may have a significant impact on the results. Scholars have drawn two very different conclusions from this observation. One view is that the experimenter should try to impose as little context as possible. The other is that the idea of context-free experiments is naïve. If the experimenter does not impose a particular context, the subjects will choose their own, leaving the

[13]Instructions are found at http://www.bi.no/users/a0111218/

experimenter even less in control (e.g., Loewenstein 1999). In accordance with the latter view, the subjects in our experiment were explicitly told in the invitation, in the general introduction, and in other instructions, that the purpose of the experiment was to test a set of hypotheses derived from a game-theoretic model that tries to identify conditions for international cooperation to curb climate change. We also said that they would not be informed about the nature of the hypotheses, explaining that this information might influence the results. This means that our findings should be interpreted with the context of climate change in mind. Without further research it is difficult to say to what extent generalizations to other contexts can be made.

*Payoffs and incentives*: All subjects received a show-up fee of NOK 300, in addition to real money made in the experiment.[14] In the experiment all costs and rewards were denominated in an experimental currency, called "schillings." The exchange rate of 0.3 NOK to a schilling was made public knowledge at the start of the experiment. The latter was done to approximate the assumption of complete information made in the benchmark model.

In the beginning of every round each subject received a certain amount of schillings (6, 12 or 21), which they could either keep or contribute in its entirety to a public good. A contribution always raised every group member's payoff by 3 schillings. However, as the minimum cost of a contribution was 6 schillings, abstaining was invariably a dominant strategy in the stage game. The unique Nash equilibrium of the stage game is therefore that all subjects keep the money for themselves. Furthermore, this equilibrium is Pareto inefficient. For example, in the low-cost case with 5 members in each group, each subject would receive 15 schillings if all group members contributed, as opposed to 6 schillings if all subjects abstained. However, the net (short-term) loss incurred by making a contribution differed significantly between treatments, depending on whether the cost of a contribution was 6, 12 or 21 schillings.

In all sessions the marginal gains of contributing and abstaining respectively were kept constant at $d = b = 3$. With a continuation probability of 0.9 we have a unique cooperative renegotiation-proof equilibrium in each cost treatment. These are as follows:

i. Exactly $m = 3$ subjects contribute in equilibrium when the cost of contributing is "low" (6 schillings).

ii. Exactly $m = 5$ subjects contribute in equilibrium when the cost of contributing is "high" (12 schillings).

iii. Exactly $m = 7$ subjects contribute in equilibrium when the cost of contributing is "very high" (21 schillings)[15].

These equilibria are exactly at an integer value, and not simply maxima as in the earlier presentation. This is due to the fact that future payoffs are discounted with a factor of 0.9 (corresponding to the continuation probability). For details concerning the restrictions on discounted payoffs, see Finus & Rundshagen (1998a:157-9). In (i) and (ii) the incentives are weak, in the sense that players are indifferent between contributing or not. In (iii) there is a positive incentive to contribute. In addition to the equilibria in (i) – (iii), playing the equilibrium of the stage game (universal defection) is always a renegotiation-proof equilibrium. The equilibria in our design give rise to a simple comparative static, from which we derive our concrete directional predictions.

*Information structure:* In the benchmark model, information is assumed to be complete and "almost perfect." In a game of complete information, the structure of the game (strategy sets and preferences) is common knowledge. This assumption is impossible to replicate in the

[14]At the time the experiment was carried out (29.04.03 and 30.04.03), 1 USD bought approximately 7.07 NOK.

[15]The last equilibrium is of course only relevant for the large group.

laboratory. However, the structure of the game was made *public* knowledge by (i) providing all subjects with payoff matrixes and other relevant information about the game, (ii) making sure that all participants could observe that all participants got this information, and (iii) using control questions to verify that the information was understood.[16] Almost perfect information implies that the history of the game up to and including round $t - 1$ must be common knowledge when subjects make their decisions in period $t$. In order to approximate this assumption in the lab, an updated statistic was generated on the screen of the subjects' computers *before* they entered the decision phase of a new round $t > 1$. This statistic contained information about the choices of every subject in every previous round of the session. Moreover, a private piece of subject specific information was revealed on the screen as a reminder: the subject's own conjecture about the number of contributors for the present round.[17] The subjects were also reminded of their own monetary payoff in the preceding round and the total number of contributions from the group in that round.

*After* entering the decision phase of a new round, all subjects had continued access to the complete history of all the subjects' choices in the session up to that round. While making decisions, subjects were also reminded of the cost of contributing. In addition, each subject had a paper copy of the relevant payoff matrix.

In the treatment including a random draw of signatories, a list of subject numbers matched with each subject's "type" ("Signatory" or "Non-signatory") was presented to the subjects *before* every round of the session. The number of signatories corresponded to the number of contributors in the renegotiation proof equilibrium (3 signatories for low costs, 5 for high costs and 7 for very high costs).

*Communication*: During the experiment, the subjects were not allowed to communicate in other ways than through their computers.[18]

*Software*: The experiment was conducted using the software Z-tree (Fischbacher 1999). Implementing the experiment in this way allowed for flexibility, while minimizing interference from the administrators of the experiment during its execution.

*Questionnaire 2*: After the final session the subjects were given a second questionnaire, prompting them to indicate, on a 0–100 scale, the emphasis they had placed on the climate dimension of the experiment relative to monetary benefits. The answers varied from 0 to 100, with an average of 35.6 and a standard deviation of 28.6.

## 5 Experimental results

We present the results in three steps. First, we look at aggregate contribution levels for all sessions, conditioned on group size, the cost of contribution, and whether subjects were explained the nature of the Getting Even equilibrium. Second, we analyse micro data in order to estimate the probability of an individual contribution, conditioned on the same variables as in the aggregate analysis, and controlling for various cross-sectional and time-dependent influences. Finally, we carry out an analysis at the level of group interaction. The dependent variable is now the average group contributions per round, conditioned on the same variables

[16] In a subsequent section we discuss a possible violation of the common knowledge assumption revealed in the experiment.

[17] We do not analyse this statistic in the present paper, but plan to utilize it in a future paper.

[18] Views differ as to whether this restriction might influence the results of the experiment. Barrett (1999) suggests that cheap talk between rounds might facilitate renegotiation-proofness. By contrast, van Damme (1989:207, note 1) claims that "[explicit communication]… is irrelevant: Even if no player can articulate the proposal, the logic underlying the argument should convince both players not to punish each other (and themselves)."

as before. This three-step approach has been chosen in order to ensure that our findings are not artefacts of aggregation, but consistent over various levels of analysis.

*Aggregate analysis:* Consider tables 1, 2a and 2b. Table 1 shows the percentage of players expected to contribute in the Getting Even renegotiation-proof equilibrium under four different combinations of group size and cost of contributing.

**Table 1: Expected contributions by group size and cost. Percentage of players expected to contribute in a renegotiation-proof equilibrium.**

| | Small group (n=5) | Large group (n=10) |
|---|---|---|
| Low cost<br>(6 schillings) | 60<br>(3 of 5) | 30<br>(3 of 10) |
| High cost<br>(12 schillings) | 100<br>(5 of 5) | 50<br>(5 of 10) |

While, according to proposition 2, the absolute number of contributors in a group should not vary if costs are kept constant, the *percentage* of players that contribute *should* vary. In particular, this percentage ought to decrease with increasing group size, for a given cost. This effect is in line with one interpretation of Olson's size principle (Olson 1965, Barrett 1999). More interesting perhaps, the percentage of players that contribute should increase with increasing costs of contributing, holding group size constant. This is the effect suggested in proposition 1.

What, then, were the results of the aggregate analysis? Consider first the case where subjects were not given any information about the nature of equilibrium play (and where no signatories therefore were drawn). The results are reported in table 2a.

**Table 2a: Contributions by group size and cost. Percent of players who contributed (N=Decisions).**

| NO SIGNATORIES DRAWN | Small group (n=5) | Large group (n=10) |
|---|---|---|
| Low cost<br>(6 schillings) | 42.7<br>(700) | 49.5<br>(400) |
| High cost<br>(12 schillings) | 17.1<br>(340) | 33.0<br>(100) |

The table clearly shows that proposition 1 is not supported by the data. To the contrary, the average percentage of contributions *falls* with increasing costs of contributions, holding group size constant. With a group of five, an average of 42.7% of the players contribute (17.3 percentage points below the contribution level expected in equilibrium) in the low-cost treatment.. Moving to the high-cost treatment, contributions fall to 17.1 percentage points on average. With large groups, contributions again fall when going from low to high costs, contrary to equilibrium expectations. In this case, the average contribution level for low cost is 49.5% (19.5 percentage points above the equilibrium expectation) and drops to 33 percentage points when moving to high cost.

In the case of large groups, we also ran sessions with “very high” costs (21 schillings, not reported in the table). For very high costs the expected equilibrium level of contributing players is 70% (7 out of 10 subjects contribute in equilibrium). However, the average contribution level observed for this treatment was only 22.7% (N=340). Thus, increasing the cost from high (12 schillings) to very high (21 schillings), reduced the average contribution level by 10.3 percentage points, discrediting proposition 1 further.

> *OBSERVATION 1A: In sessions carried out before subjects were explained the nature of equilibrium play – and given clues as to which subjects were expected to contribute in equilibrium – average contribution levels fell markedly with rising costs.*

Although the Getting Even strategy admits only one cooperative renegotiation-proof equilibrium in each of our experimental controls, this equilibrium can usually be achieved by many different combinations of specific subjects as signatories and non-signatories.[19] A possible explanation of obervation 1a is that the subjects were simply unable to coordinate on one particular of the (often) bewilderingly large number of ways to play the equilibrium. We therefore conducted a number of sessions in which the subjects (i) were explained the nature of the Getting Even renegotiation proof equilibrium, *and* (ii) an appropriate number of “signatories” were randomly selected at the beginning of each session. In these sessions, the administrators *de facto* solved the coordination problem faced by the subjects. The results from these sessions are reported in table 2b.

**Table 2b: Contributions by group size and cost. Percent of players who contributed (N=Decisions).**

| SIGNATORIES DRAWN | Small group (n=5) | Large group (n=10) |
|---|---|---|
| Low cost<br>(6 schillings) | 54.1<br>(1060) | 52.4<br>(700) |
| High cost<br>(12 schillings) | 16.4<br>(360) | 29.0<br>(780) |

It is evident from the table that proposition 1 does not hold *even* when the subjects are explained the nature of the renegotiation-proof equilibrium and the administrators single out the subjects expected to contribute in equilibrium. Again, the average number of players contributing *decreases* with increasing costs of contributions, holding group size constant. In

[19]With low cost there exists – by the binomial formula – 10 ways to play the equilibrium in the small group. With low cost and a large group, the number of ways to play the equilibrium increases to 120. With high cost, there is one way to play the equilibrium in the small group, while the number of ways to play the equilibrium in the large group is 252. If costs are very high (only relevant for the large group), the number of ways to play the equilibrium is again 120. Since it is more profitable to be a non-contributor than a contributor in equilibrium, the subjects face formidable coordination problems in trying to find a way to play the renegotiation-proof equilibrium. The exception is the high-cost, small-group treatment. While in this treatment average contribution levels fall short of the equilibrium expectation by far, we do not hold this observation against the theory, since it builds on a point and not a directional hypothesis. Inability to solve the coordination problems may of course destroy our directional predictions.

small groups, while an average of 54.1% of the players contributed when costs were low,[20] only 16.4% contributed when costs were high. In large groups, the contribution level was 52.4% with low costs, and 29.0% with high costs. In short, increasing the cost of making a contribution causes the the overall level of cooperation to drop.

In the large group condition, we ran additional sessions with very high costs (21 schillings). The average contribution level observed for this treatment was only 23.8% (N=240). Thus, increasing the cost from high (12 schillings) to very high (21 schillings) further reduced the average contribution level by *5.2* percentage points. The results can be summed up in the following observation:

> *OBSERVATION 1B: After subjects were told about the renegotiation-proof way to play the game – and given clues as to which subjects were expected to contribute in equilibrium – average contribution levels still fell markedly with rising costs.*

What about proposition 2? Observing table 2a and 2b, we see that the average level of contributions tends to increase when moving from small to large groups, holding cost levels constant. In the low-cost treatment, however, this effect is somewhat ambiguous. If no signatories are drawn (table 2a), the contribution level increases by 6.8 percentage points when moving from small to large groups. If signatories are drawn (table 2b), contribution levels decrease by 1.7 percentage points when moving from small to large groups. When costs are high, the effect is more pronounced: If no signatories are drawn, moving from the small to the large group increases the average level of contribution by 15.9 percentage points. The comparable number when signatories are drawn is an increase of 12.6 percentage points. This pattern leads to the following observation:

> *OBSERVATION 2: Moving from small to large groups increases the average number of contributions markedly if costs are high. If costs are low, the effect of increasing group size is ambiguous.*

What is clear from observation 2 is that the independence conjectured in proposition 2 is not observed for large groups. Increasing group size has an effect on the average level of contributions, and the effect is strong and positive for high costs, implying an interaction between costs and group size. The tendency observed (at least for high costs) is the *opposite* of the size effect predicted by Olson (1965).

How can the observed pattern be explained? In our design the marginal per capita return is 0.6 in the small group treatment, as compared to 0.3 in the large group treatment. Other things being equal, one should therefore – based on previous experiments – expect average contributions to be smaller in a large group than in a small group (Ledyard 1995:137-41;149-55, Isaac, Walker and Thomas 1984). However, we observe an effect in the opposite direction. Thus, we suspect that group size has an independent effect, and this effect goes in the opposite direction of the one predicted by standard public goods models. In Barrett's model, increasing group size should not lead to changes in observed behaviour. By contrast, the data suggests that group size matters.

---

[20]It may be noticed that this is 5.9 percentage points below the contribution level expected in equilibrium.

For all of our 10 experimental treatments, a non-parametric test (based on the binomial distribution) was carried out. The test was designed to check whether the observed distribution of contributions differed significantly from the one expected in the renegotiation-proof equilibrium. In all 10 treatments, we tested the null hypothesis that the actual distributions correspond to the distribution expected in the renegotiation-proof equilibria.[21] In every case, the null hypothesis was rejected at significance levels down to (and below) 0.001. This leads to our third observation, which could be seen as an illustration of how easy it is to refute a point prediction (our null):

> *OBSERVATION 3*: *The distribution of contributions is significantly different from the ones one would expect to see if the players had played the Getting Even renegotiation-proof equilibrium.*

*Analysis of micro data:* We now turn to the results of our micro level analysis. Table 3 shows the estimates of four logistical regressions. The first two equations are of the fixed effect type. This means that cross-sectional (subject specific and inter-group specific variation) and time-specific (session numbers and round numbers) variance have been eliminated. What we want to estimate is the probability that a subject will contribute, given group size and cost of contribution. Both equations include the variables GROUP SIZE, HIGH COST and VERY HIGH COST. The first of these variables is a dummy taking the value one if the group is large (N=10), and zero otherwise. The second variable is a dummy that takes the value one if costs are high (12 schillings), and zero otherwise. The third variable is a dummy scoring one if costs are very high (21 schillings), and zero otherwise. The first equation (model 1) contains only cases where subjects have been told how to play the renegotiation-proof equilibrium. The variable TYPE takes the value one if the subject was singled out as a signatory, and zero otherwise. The second equation (model 2) contains only cases where subjects were *not* told how to play the renegotiation proof equilibrium. In this equation the variable TYPE has no meaning, since no random draw of signatories was carried out.

As can be seen from model 1 and model 2, the regression coefficients for group size and cost dummies are highly significant and have signs in disagreement with propositions 1 and 2, controlling for other cross-sectional and time-specific variance.[22] Going from a small to a large group significantly increases the probability of an individual contribution. Furthermore, moving from low to high cost significantly decreases the probability of an individual contribution. Lastly, going from low to very high cost also significantly decreases the probability of an individual contribution. Regarding model 1, we see that solving the coordination problem for the subjects – by instructing them on how to play the renegotiation-proof equilibrium and drawing the appropriate number of signatories – does not alter these conclusions. This aside, being drawn as a signatory significantly increases the probability of contributing to the collective good. As is evident from the probability curves presented in figure 2, however, this effect is miniscule. This is quite remarkable, since in this experimental treatment Barrett's model is given considerable help. Nevertheless, the results offer very little support for proposition 1.

[21]The expected distributions in the renegotiation-proof equilibria are given in table 1 for low and high costs, and in the text for very high costs.

[22] See Köningstein 1998 for details in fixed effects models and experimental data.

**Table 3: Dependent variable: individual decision (contribution=1). Logistical regression coefficients (standard errors).**

| | Model 1 † | Model 2 † | Model 3 | Model 4 |
|---|---|---|---|---|
| CONSTANT | 2.37 (1.60) | -0.86 (.644) | -2.13*** (.143) | -1.66*** (.140) |
| GROUP SIZE | 1.79** (.613) | 2.05*** (.626) | 0.55** (.102) | 0.48*** (.101) |
| HIGH COST | -1.73** (.259) | -1.48*** (.361) | -1.85*** (.112) | -1.86*** (.112) |
| VERY HIGH COST | -1.29* (.534) | -1.94*** (.483) | -2.11*** (.209) | -2.13*** (.208) |
| TYPE | 0.71*** (.161) | _ | 0.44*** (.103) | _ |
| CLIMATE WEIGHT | _ | _ | 4.10*** (.193) | 4.31*** (.191) |
| GENDER | _ | _ | 0.66*** (.100) | 0.73*** (.099) |
| Correctly predicted (%) | 79.8 | 77.8 | 76.3 | 68.4 |
| Model Chi-square | 1611.8*** | 771.8*** | 1010.1*** | 312.4*** |
| N | 3140 | 1880 | 3140 | 1880 |

*** $p < 0.001$

** $p < 0.05$

* $p < 0.10$

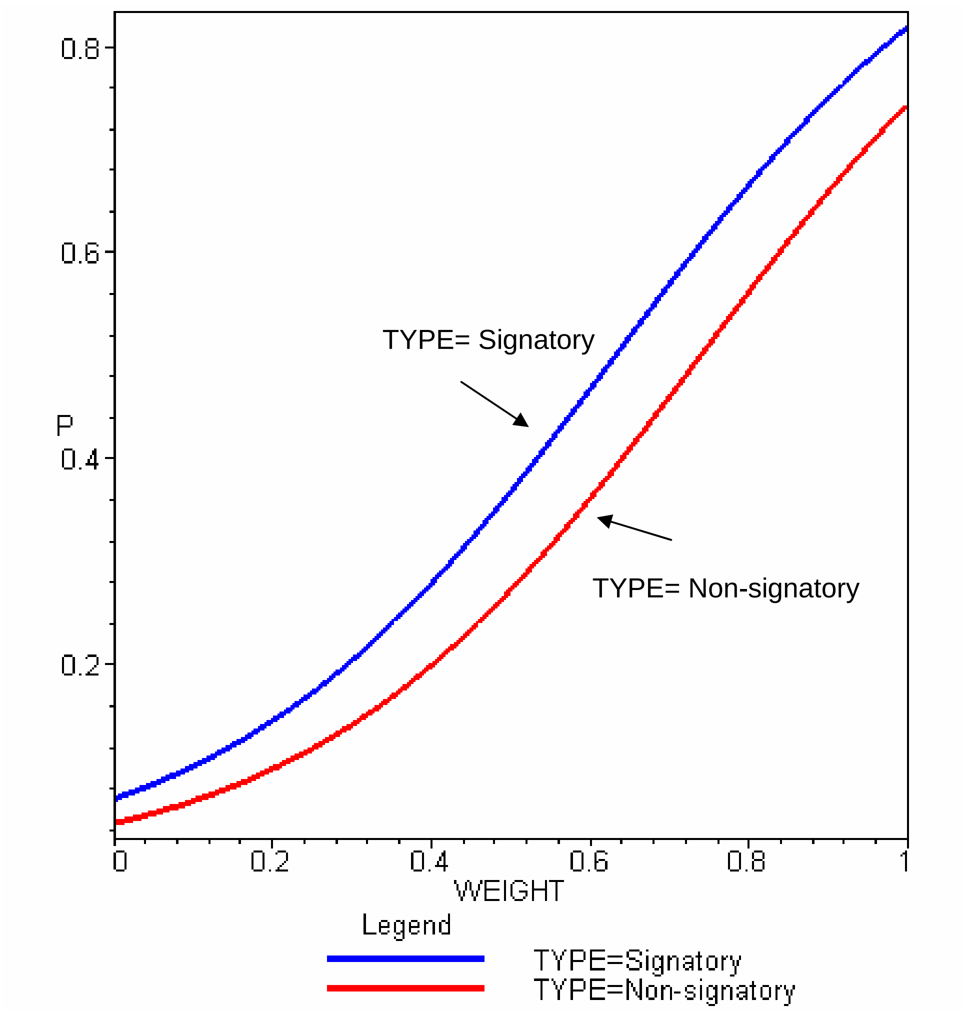
† Fixed effects model: Controls for subject dummies, group dummies, round numbers and session numbers (coefficients not reported)

Let us now turn to models 3 and 4. These are not fixed effects models. Instead we have included two subject-specific variables. GENDER takes the value of one if the subject is a male, and zero otherwise. CLIMATE WEIGHT is a continuous variable based on questionnaire 2, and rescaled to the interval [0, 1]. This variable, then, takes a higher value, the more weight the subject says he or she placed on climate (as compared to monetary rewards). Model 3 contains only cases where subjects were explained the nature of the Getting Even equilibrium, and where “signatories” were appointed. Model 4 contains only cases where subjects were not explained the nature of the equilibrium.

The results concerning the dummies for group size and cost level are by and large as in the fixed effects models. The same is true for the TYPE DRAW dummy in model 3. The results for GENDER show that male subjects are significantly more likely than females to contribute to the collective good. The result is robust over the conditioning on equilibrium story and type draw. The sign of this coefficient is in line with a number of studies, but at odds with some others (cf. Ledyard 1995:160-1).

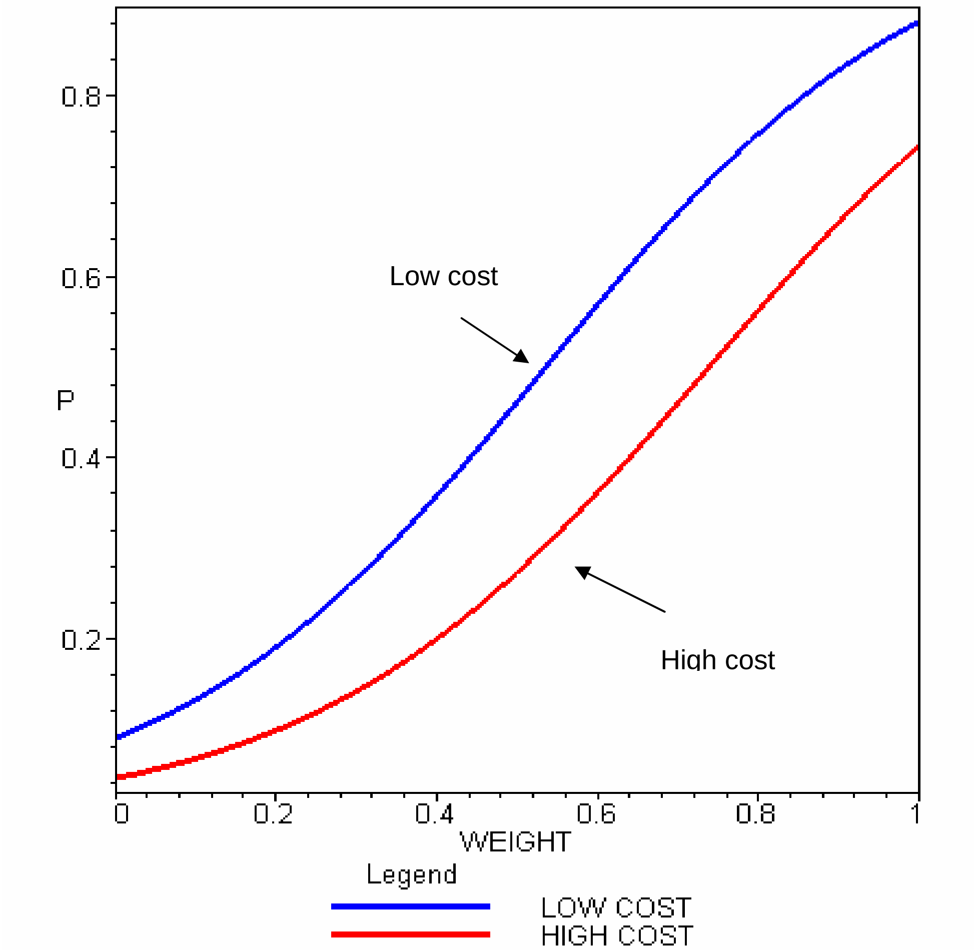
Figures 2 and 3 offer comparable probability curves for selected variable values. Figure 2 is based on the regression in model 3. We look at the probability that a male subject will contribute in a large group when costs are very high. In the upper curve, a male was singled out as a signatory, in the lower curve he was singled out as a non signatory. Both curves are drawn as functions of the CLIMATE WEIGHT.

**Figure 2: Probability curves based on model 3 (table 3). Large group, male, very high costs.**




In figure 2 the difference in the probability of contributing never exceeds 7 percentage points. That is, the probability of a contribution from a male signatory facing very high costs in a large group interaction is never (for no value of the CLIMATE INDEX) more than 7 percentage points higher than the probability of a contribution from a male non-signatory facing the same cost in the same sized group interaction.

**Figure 3: Probability curves based on model 4 (table 3). Small group, male.**




In figure 3 the difference in the probability of contributing reaches a maximum of 20 percent points. That is, the probability of a contribution from a male facing high costs in a small group is at most 20 percentage points lower than the probability of a contribution from a male facing low costs in the same sized group interaction.

A remarkable finding revealed by the probability curves in figures 2 and 3 is the potency of the climate weight: The behaviour of the subjects seems to a large extent to be conditioned on their "climate preferences." There is something deeply disturbing in this finding, since nothing subjects did in the experiment could possibly have any effect on the climate. Why, then, condition one's behaviour on one's attitude towards climate policy? Heterogeneity in the subjects' stated climate preferences seems to drive the experimental results to a high degree.

One possible way to interpret this finding is that the subjects differed with respect to their "social preferences," i.e., the extent to which they care about other goals than their own monetary payoffs. However, most notions of social preferences assume some form of instrumental – though not purely self interested – goal seeking (Bolton & Ockenfels 2000, Camerer 2003, Charness & Rabin 2001, Fehr & Schmidt 1999).[23] In our view, explaining the behaviour in question with a reference to social preferences strains this concept too far.

In any case, the climate weight is an important predictor of individual tendencies to contribute. This suggests that individual preferences do not necessarily correspond to the induced preferences given by the payoff matrixes, which in turn challenges an important assumption in Barrett's model – that individual preferences are common knowledge. Violating this assumption could vitally influence the equilibria in the experimental game.

That said, however, controlling for variation in the stated climate preferences does not change the main conclusions. We still find that the probability of a contribution is significantly reduced with higher costs but increased with larger group size. Thus, propositions 1 and 2 do not gain support when controlled for the effect of variation in stated climate preferences. Perhaps more surprisingly, giving Barrett's model a considerable amount of "help" – by explaining to the subjects how to play his suggested equilibrium and solving the coordination problems involved – does not produce additional support for his central model prediction (proposition 1). The analysis of micro data warrants the following observations:

> *OBSERVATION 4: Other things being equal, a subject is less likely to contribute the higher the cost of a contribution.*
>
> *OBSERVATION 5: Other things being equal, a subject is more likely to contribute in a large group than in a small group.*
>
> *OBSERVATION 6: Other things being equal, a male subject is more likely to contribute than a female subject.*
>
> *OBSERVATION 7: Other things being equal, a subject is more likely to contribute the greater the weight the subject attaches to the climate relative to pecuniary gains.*

*Analysis of group interactions:* We now focus on the unit of strategic interaction, that is, the group in which the subjects actually interact. Table 4 presents the results of the analysis.

The dependent variable in table 4 is the average percentage of players that contributed in each round. As independent variables we use the cost-dummies and the group-size dummy, as defined above. To account for dynamics, we control for round number. The dummy variable TYPE DRAW takes the value of one if subjects were told how to play and signatories were drawn, and zero otherwise. Since a Spearman rank correlation test indicated signs of heteroscedasticity, a generalized least squares regression (GLS) was run in addition to ordinary least squares (OLS).[24] As can be seen, the difference between the two is marginal.

The results point in the same direction as our previous analysis. Moving one step up the cost ladder reduces the average fraction of players that contributed by roughly 20 to 30 percentage points. Moving from small to large groups increases the average percentage of players that contributed by roughly 5 to 10 percentage points. Being explained the nature of the Getting Even equilibrium leads to an increase of contributions of roughly the same magnitude. No dynamic effects are discernible.

---

[23]An exception is the non-instrumental behaviour found in Fehr & Gächter (2000). A notable fraction of the subjects in the public goods experiment studied punished players that did not contribute, even though punishment was costly and could have no effect whatsoever on the punisher's future payoffs.

[24]Details of the Spearman test are found in Gujarati 1995:372-4.

**Table 4: Dependent variable: average group contributions per round. Unstandardized regression -coefficients (standard errors).**

| | OLS | GLS |
|---|---|---|
| Constant | 0.44*** (.019) | 0.58*** (.010) |
| ROUND NUMBER | 0.00 (.001) | 0.00 (.001) |
| GROUP SIZE | 0.05*** (.022) | 0.08*** (.010) |
| HIGH COST | -0.29*** (.022) | -0.31*** (.011) |
| VERY HIGH COST | -0.29*** (.041) | -0.23*** (.022) |
| TYPE DRAW | 0.06*** (.022) | 0.12*** (.010) |
| F-statistics | 45.3*** | 187.0*** |
| Adjusted $R^2$ | 0.23 | 0.73 |
| N | 748 | 748 |

*** $p < 0.001$

** $p < 0.050$

* $p < 0.100$

The latter point is worth elaborating on. In PD games with a fixed horizon, a robust experimental finding is that contributions tend to decrease over time, and that free riding is especially pronounced in the last few rounds.[25] In our data, there is no tendency of more pronounced free riding towards the end of each session. In fact, there is no particular dynamic pattern at all. The reason may well be precisely that the time horizon is not fixed.

Figures 4 and 5 exemplify this point. The figures display developments over time for the longest session in the small group treatment (Figure 4), and the longest session in the large group treatment (Figure 5). The former session ended after 48 rounds, the latter after 37 rounds. In the experimental literature, conclusions about convergence are in general drawn on the basis of observed behaviour over 10 to 20 rounds in fixed horizon games. The sessions shown in figures 4 and 5 were considerably longer than this, but convergence is still absent. More interestingly, these sessions are representative, in the sense that no convergence can be found in any of the other sessions in the experiment either.

[25]See Fehr and Schmidt 1999 for a summary of such results.

**Figure 4: Average fraction of subjects that contribute. Session 11, 48 rounds. Cost 6 schillings, no type draw. Groups 1, 2, 3 and 4 from left to right on x-axis. The dotted line represents the equilibrium expectation.**

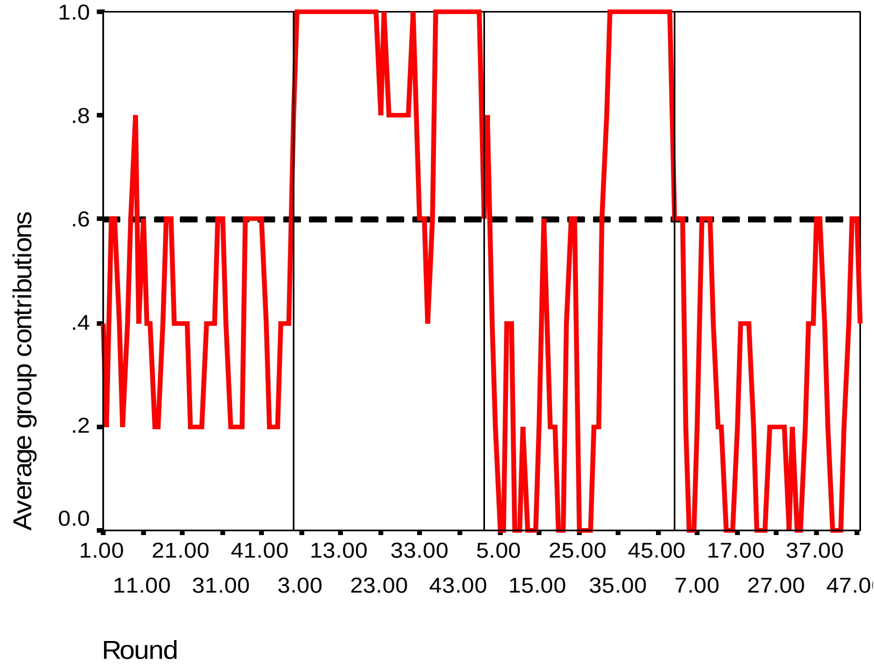
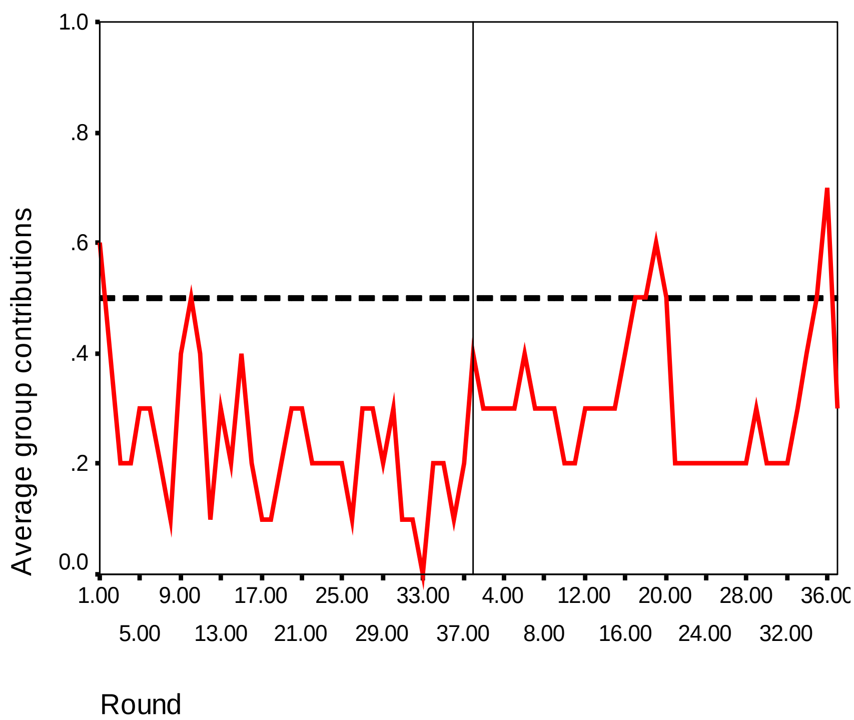



**Figure 5: Average group contributions. Large groups. Session 23, 37 rounds. Cost 12 schillings, types drawn. Groups 1 and 2 from left to right on x-axis. The dotted line represents the equilibrium expectation.**



From the analysis of group interactions we may explicate the following observations:

> *OBSERVATION 8: Other things being equal, increasing the cost of a contribution reduces the average fraction of group members that contribute per round.*
>
> *OBSERVATION 9: Other things being equal, increasing group size increases the average fraction of group members that contribute per round. This effect is moderate.*
>
> *OBSERVATION 10: Group behaviour does not change systematically over time.*
>
> *OBSERVATION 11: Explaining the nature of the Getting Even renegotiation-proof equilibrium to the subjects, and solving the coordination problem for the group, increases the average fraction of group members that contribute per round, other things being equal. This effect is moderate.*

# 6 Conclusions

The notion of a renegotiation-proof equilibrium has become central to game-theoretical thinking about international environmental agreements in general and climate agreements in particular. However, requiring renegotiation-proofness leads to very pessimistic predictions in contexts such as the repeated N-person Prisoners' Dilemma. Also, some of these predictions are highly counterintuitive. In particular, it strains common sense that higher costs should induce *more* cooperation and that the number of contributors should be independent of group size. This paper has reported the results of an experiment that was carefully designed to test these two predictions. The main findings are that high costs have a strong *negative* effect, and group size a marked (although weaker) *positive* effect on cooperation.

As noted above, Barrett's work has (rightly) had a profound influence on the literature on international environmental cooperation. One of the reasons for this is that Barrett has been able to produce an impressive number of interesting – and sometimes provocative – hypotheses. However, at least some of these hypotheses seem to be grounded on very strong and indeed bold assumptions. It is therefore important to confront these hypotheses with empirical evidence. This is what we have done in this paper. Needless to say, there is a long way from individual behavior in the laboratory to governmental decisions on the international scene. Thus, the results of experiments such as the one reported in this paper should be interpreted with care. Yet we believe that one should take seriously the fact that our findings are not easily reconciled with the predictions generated by the repeated Prisoners' Dilemma model. At best, important questions remain unanswered. At worst, our results challenge the empirical relevance of this model, which underlies a good deal of recent theorizing on international climate agreements.

## References

Banks, Jeffrey, Colin Camerer and David Porter. 1994. "Experimental tests of Nash refinements in signaling games." Games and Economic Behavior 6:1-31.

Barrett, Scott. 2003. Environment & Statecraft. Oxford: Oxford University Press.

Barrett, Scott. 2002. "Consensus Treaties." Journal of Institutional and Theoretical Economics 4:529-47.

Barrett, Scott. 2001. "International Cooperation for Sale." European Economic Review 45:1835-50.

Barrett, Scott 1999a. "A Theory of Full International Cooperation." Journal of Theoretical Politics 11:519-41.

Barrett, Scott 1999b. "Political Economy of the Kyoto Protocol." Oxford Economic Papers 14:20-39.

Barrett, Scott. 1994. "Self-enforcing International Environmental Agreements." Oxford Economic Papers 46:878-94.

Bergin, James and Bentley W. MacLeod. 1993. "Efficiency and Renegotiation in Repeated Games." Journal of Economic Theory 61:42-73.

Brants, Jordi and Charles Plott. 1995. "Limitations of dominance and forward induction: Experimental evidence." Economic Letters 49:391-95.

Brants, Jordi and Charles Plott. 1993. "Adjustment patterns and equilibrium selection in experimental signaling games". International Journal of Game Theory 22:279-302.

Brants, Jordi and Charles Plott. 1992. "An experimental test of equilibrium dominance in signaling games." American Economic Review 82:1350-65.

Camerer, Colin. 2003. Behavioral game theory. Experiments in strategic interaction. Princeton: Princeton University Press.

Carraro, Carlo. 1999. "The Structure of International Environmental Agreements." In International Environmental Agreements on Climate Change, ed. Carlo Carraro. London: Kluwer Academic Press.

Carraro, Carlo and Domenico Siniscalco. 1998. "International Environmental Agreements: Incentives and Political Economy." European Economic Review 42:561-72.

Chamberlin, John R. 1982. "Provision of Public Goods as a Function of Group Size." In Rational Man and Irrational Society, eds. Brian Barry and Russell Hardin. Beverly Hills: Sage.

Dawes, Robyn and Richard Thaler. 1988. "Cooperation." Journal of Economic Perspectives 2(3):187-97.

Ecchia, Giulio and Marco Mariotti. 1998. "Coalition Formation in International Agreements and the Role of Institutions." European Economic Review 42:573-82.

Enders, Alfred and Michael Finus. 2002. "Quotas Beat Taxes in a Global Emission Game." International Tax and Public Finance 9(6): 687-707.

Farrell, James. 1983. "Credible Repeated Game Equilibrium." Unpublished manuscript.

Farell, James and Eric Maskin. 1989. "Renegotiation in Repeated Games." Games and Economic Behavior 1:327-60.

Fehr, Ernst and Simon Gächter. 2000. "Cooperation and Punishment in Public Goods Experiments." Journal of Economic Perspectives 14: 159-81.

Fehr, Ernst and Klaus M. Schmidt. 1999. "A Theory of Fairness, Competition and Cooperation." Journal of Economics 114(3):817-68.

Finus, Michael. 2001. Game Theory and International Environmental Cooperation. Cheltenham: Edward Elgar.

Finus, Michael and Bianca Rundshagen. 1998a. "Toward a Positive Theory of Coalition Formation and Endogenous Choice in Global Pollution Control." Public Choice 96:145-86.

Finus, Michael and Bianca Rundshagen. 1998b. "Renegotiation-proof Equilibria in a Global Emission Game when Players Are Impatient. Environmental and Resource Economics 12(3):275-306.

Fischbacher, Urs. 1999. z-Tree: Zurich Toolbox for Readymade Economic Experiments. Mimeo. University of Zurich. (http://www.unizh.ch/iew/wp/iewwp021.pdf)

Fredriksson, Per G. and Noel Gaston. 2000. "Ratification of the 1992 climate change convention: What determines legislative delay?" Public Choice 104:345-68.

Gächter, Simon and Ernst Fehr. 1999. "Collective action as a social exchange." Journal of Economic Behavior and Organization 39:342-69.

Gujarati, Damodar N. 1995. Basic Econometrics (3rd ed.). London: McGraw Hill.

Hovi Jon, Detlef Sprinz and Arild Underdal. 2003. "The Oslo-Potsdam Solution to Measuring Regime Effectiveness: Critique, Response, and the Road Ahead." Global Environmental Politics 3(3):105-107.

Isaac, Mark R., James M. Walker and Susan H. Thomas. 1984. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." Public Choice 43:113-49.

Isaac, Mark R., Kenneth F. McCue and Charles R. Plott. 1985. "Public goods provision in an experimental environment." Journal of Public Economics 26:51-74.

Köningstein, Manfred. 1998. "Measuring Treatement Effects in Experimental Cross-Sectional Time Series." Research paper. Berlin: Humbolt University. (http://www.wiwi.hu-berlin.de/wt3/Forschung/Veroffentlichungen/WorkPapers/workpapers.html)

Ledyard, John O. 1995. "Public Goods: A Survey of Experimental Research." In Handbook of Experimental Economics, eds. James Kagel and Alvin Roth. Princeton: Princeton University Press.

Loewenstein, George. 1999. "Experimental Economics from the Vantage-Point of Behavioural Economics." The Economic Journal 109 (February):F25-F34.

Marwell, Gerald and R. Ames. 1979. "Experiments on the Provision of Public Goods I: Resources, Interest, Group Size, and the Free Rider Problem. American Journal of Sociology 84:1335-60.

Olson, Mancur. 1965. The Logic of Collective Action: Public Goods and the Theory of Groups. Cambridge Mass.: Harvard University Press.

Partow, Zeinab and Andrew Scotter. 1996. Does game theory predict well for the wrong reasons? An experimental investigation. New York University: Starr Center for Applied Economics. Working Paper 93-46.

Potters, Jan and Franz van Winden. 2000. "Professionals and students in a lobbying experiment." Journal of Economic Behavior and Organization 43:499-522.

Potters, Jan and Franz van Winden. 1996. "Comparative statics of a signaling game: An experimental study." International Journal of Game Theory 25:329-53.

Roth, Alvin and John Murnighan. 1978. "Equilibrium Behavior and Repeated Play of the Prisoners Dilemma." Journal of Mathematical Psychology 17:189-98.

Samuelson, Larry. 1997. Evolutionary Games and Equilibrium Selection. Cambridge Mass.: The MIT-Press.

Schelling, Thomas. 1992. "Some Economics of Global Warming." American Economic Review 82 [1]:1-14.

UNEP 2001. Climate Change 2001. Synthesis Report. Contribution of Working Groups I, II and III to the Third Assessment Report of the Intergovernmental Panel on Climate Change. Cambridge: Cambridge University Press.

Van Damme, Eric. 1989. Renegotiation-proof Equilibria in Repeated Prisoners' Dilemma. Journal of Economic Theory 47:206-17.

Wagner, Ulrich. 2001. "The Design of Stable International Environmental Agreements: Economic Theory and Political Economy." Journal of Economic Surveys 15:377-411.