Abstract:
This paper uses three different experimental treatments – a standard two-player Prisoner’s Dilemma, a direct democracy, and a representative democracy – to analyze the effects of grouping on individual decision-making. A major finding is that a subject’s propensity to cooperate is much greater when she is embedded in a group compared to when she acts individually. Further, in the representative democracy treatments, cooperation occurs in latter periods of the experiment, a result that is novel to the Prisoner’s Dilemma literature. These results suggest that the theoretical literature that treats groups as equivalent to unitary rational actors is missing an important link in the intergroup decision-making process.

Keywords: experiments, voting, international environmental problems, Prisoner’s Dilemma

JEL: C91, C92, D72, Q28

*Corresponding author: Stephan Kroll, Department of Economics & Finance, University of Wyoming, Laramie, WY 82071-3985. Phone: (307) 766-6240; Fax: (307) 766-5090; e-mail: Kroll@uwyo.edu
International Environmental Problems as Two-Level Games:
An Experimental Investigation

1. Introduction

Standard economic models of international relations often use the paradigm of a country, or its
government, as a unitary rational actor that has only the welfare of its population in mind. Recently, the
literature on “two-level games” has evolved, which pays tribute to domestic constraints on international
decision-making. The underlying presumption in the models of this literature is that a government plays (at
least) two games simultaneously, one with its counterparts in other countries and one with its own electorate.
A striking feature of the links between these two games is that a decision can be rational in one game but
irrational in the other.

In this paper we report results from experiments on two-level games. Using three different
treatments, we compare and contrast outcomes under individual games with outcomes when decision-makers
are embedded in a group. We model two different types of group games – a form of direct democracy where
members vote directly on the group decision, and a representative democracy where voters elect a
representative who makes the decision for his or her group. A primary result from our laboratory experiments
is that individuals in groups are more inclined to cooperate than subjects making individual decisions.
Moreover, we observe groups in representative democracies cooperating even in the final periods of the
experiment. These findings have consequences for the theoretical literature of international environmental
cooperation since this literature predominantly assumes governments are unitary actors. They suggest that
adding domestic components can mitigate the rather disillusioning findings concerning cooperation between
countries.

The balance of the paper is organized as follows: In the next section, we introduce the basic model
and review related experiments from the public good, voting, and social psychology literature. Section 3
presents the experimental set-up. Section 4 reports results, and Section 5 concludes and provides natural
extensions of our work.

2. Theoretical and Experimental Background

2.1 International Environmental Problems and the Prisoners’ Dilemma

Environmental problems are often international in nature – i.e., transboundary emissions, ozone
depletion, and the Greenhouse effect. The basic game theoretic model to derive solutions to such problems
assumes that a country benefits from the aggregate effort, however its own welfare is a decreasing function in
its own contribution. Thus, countries face a social dilemma – each country prefers to contribute less than
what is optimal for the community of countries (free-ridership). For example, assume the environmental problem in question is carbon emissions (the primary precursor to global warming), and the effort of each country is pollution abatement. As such, the welfare of country $i$ can be described by

$$W_i(q_1, q_2, \ldots q_n) = B_i(q_1, q_2, \ldots q_n) - C_i(q_i),$$

where $Q = (q_1, q_2, \ldots q_n)$ is the vector of abatement measures of the $n$ countries involved, $B_i$ is the benefit for country $i$ as a function of the aggregate abatement measures, and $C_i$ is the cost for country $i$ from $q_i$.\(^1\) For expositional purposes, assume benefit and cost functions take the form

$$B_i(Q) = \omega_i Q$$

and

$$C_i(q_i) = \frac{1}{2} c_i q_i^2$$

where $\omega_i$ and $c_i$ are country-specific benefit and cost parameters (Barrett, 1997). In the (one-shot) Nash-equilibrium of this game, all countries maximize their own welfare irrespective of other countries. Hence, they choose

$$q_{i*} = \max_{q_i} (B_i(Q) - C_i(q_i)),$$

which leads to

$$q_{i*} = \omega_i/c_i.$$

If instead all countries are willing to cooperate, an equally weighted cooperative solution $q_i^c$ stems from

$$\max_{q_1, q_2, \ldots q_n} \sum_i (B_i - C_i),$$

which yields

$$q_i^c = (\omega_1 + \omega_2 + \ldots + \omega_n)/c_i.$$

\(^1\) An equivalent model takes the level of emissions instead of the level of abatement as the independent variable.
Although this simple model appeals to intuition as it reveals the non-optimal outcomes typical of the unitary-actor prisoners’ dilemma game, we do not focus on the country-level interaction, but rather we investigate the effects of placing subjects in groups. Consequently, we opt for the most simplistic form of the game by setting $n = 2$ and $\alpha_i = c_i = 1$. These assumptions yield $q_i^* = 1$ and $q_i^c = 2$, and the following payoff matrix:

(Table 1 here)

Table 1 represents the standard Prisoners’ Dilemma game. The welfare for each nation is $U^* = 1.5$ if both nations choose the Nash-equilibrium (non-cooperation), and $U^C = 2$ if both cooperate. Yet, if one nation cooperates and chooses the higher level of abatement while the other does not cooperate, the welfare levels are $U = 2.5$ for the non-cooperating nation and $U = 1$ for the cooperating nation. Experiments analyzing the behavior of individual players in this simple Prisoners’ Dilemma measure in the hundreds. The results of these experiments overwhelmingly imply that most subjects do not play dominant strategies (“free-riding” in public goods experiments, “not cooperating” in Prisoner’s Dilemmas), but their willingness to behave in the interests of the group declines in latter periods.

2.2 Voting Behavior

Although we focus on environmental issues, our experiment touches on two topics of contemporary interest for a wider audience in public choice: direct democracy versus representative democracy, and retrospective versus perspective voting.

- Two recent articles tackled the question of the optimal government from different perspectives: Niskanen (1997) examines fiscal decisions of an autocratic and a democratic government and finds that output of an economy with a democratic government is generally higher, but depends largely on the fiscal horizon of those who govern, a horizon that can be shortened by frequent elections. Kollman et al. (1997) compare direct democracy, winner-takes-all representative democracy, and proportional representation in a computer simulation. They show that the performance of these political institutions depends largely on whether citizens live in a single jurisdiction or in multiple jurisdictions. Direct democracy performs better in a single jurisdiction, representative democracies perform better in multiple jurisdictions. In a related vein, our experiment approaches the question of the optimal government from a different angle – when

---

2 A recent experiment by Nalbantian and Schotter (1997) investigates public good behavior in various institutional settings. Even though their motivation differs from ours (their goal is to find the best group incentive program to foster the productivity of workers), the basic set-up is similar. Also, see Ledyard (1995) for an excellent overview of public good experiments with more than two subjects in a group.
two groups (or jurisdictions) are in a dilemma situation, which form of democracy is better suited to solve this dilemma?

- The voting game we use in the Representative Democracy treatment follows Williams (1994), who examines if voters prefer to use a traditional purely retrospective voting rule or a retrospective-prospective rule. Using the former rule, the electorate vote for the incumbent when their payoffs in the last period are higher than in the next-to-last period; using the latter rule, voters compare the last payoff they received from the incumbent with the payoff they expect to receive in the next period from the challenger; where their expectations are shaped by the challenger’s non-binding announcement. Williams finds that in times of stability voters tend to use the traditional voting rule.

2.3 Two-Level Games

The term “two-level games” was coined by Putnam (1988), who investigated behavior of politicians at the G7-summit in 1977. Putnam’s (1988) basic intuition is that negotiators on the international level (level 1) simultaneously play another game at the domestic level (level 2). The aggregated preferences on level 2 constrain the bargaining-set on level 1. These constraints, however, are potentially advantageous for the negotiator since he/she can improve his/her relative position by pointing out that certain decisions on level 1 that are deemed rationale for a unitary actor are not feasible for a representative since they are unacceptable on level 2. This idea stems from the assumption that the preferences of the domestic actors differ from those of the representative, or at least are unknown. In our experiments laid out below, we do not make this assumption – voters and political candidates in one group have the same utility functions with regard to the international issue at hand.

To our knowledge, laboratory investigations of two-level games have not been conducted yet. Nevertheless, studies closest in spirit to two-level experiments can be found in the recent literature in social psychology, which reports results from experiments in team or group games (see, for example, Bornstein, 1992, Bornstein and Ben-Yossef, 1994, Bornstein et al., 1996). In these experiments, however, the Prisoner’s Dilemma does not appear on level 1, the intergroup-level, but on level 2, the intragroup-level. Despite the somewhat confusing name “Intergroup Prisoner’s Dilemma (IPD),” Bornstein and associates model level 1 as a zero-sum game, in which a win for one group is automatically a loss for the other group. While Bornstein and Ben-Yossef (1994) model the intragroup-decisions as an one-shot Prisoner’s Dilemma, in which group members can only choose between contribute or not, Bornstein et al. (1996) design level 2 as a repeated public-good game, where group members have discretion on how much they want to contribute.

---

Both experiments confirm what social psychologists have previously hypothesized – people are more willing to sacrifice self-interests for group causes when the social dilemma is embedded in an intergroup conflict.

3. Experimental Design

The subjects in our experiments consisted of 58 undergraduate students in five sessions at the University of Central Florida and 62 undergraduate students in five sessions at the University of Wyoming. Each subject took part in one of three treatments. Each of the ten sessions lasted approximately one hour, and subjects earned between $10 and $25. None of the subjects had previously taken part in a laboratory Prisoners’ Dilemma experiment, and only a few had been exposed to any lectures on game theory. Each session lasted 25 periods, a number that was common knowledge to the subjects. For each treatment, we used a payoff table similar to Table 1. To suit experimental needs, however, we changed the payoffs slightly by using a utility function $U = 20(B - C) - 15$, whereby $B$ and $C$ are the benefit and cost functions, respectively, from Section 2.1. Hence, the new payoff table is represented by

(Table 2 here)

The numbers in the cells represent a fictitious currency called tokens, where 40 tokens equaled $1.

In the experiments conducted in Central Florida each subject had payoff tables and strategy sheets in front of him or her. In each period, a subject had to fill out the strategy sheet and hand it to the monitor, who calculated the payoffs and gave the sheets back to the subjects. In Wyoming, subjects made their decisions on linked computers. In all treatments, subjects did not know who was in their group or against whom they played. Since our experiments were primarily designed to test whether subjects behave differently when they are embedded in a group game instead of acting on their own behalf, we ran one baseline treatment with individual decisions and two treatments with group games. Consider each experiment in more detail:

---

4 Nalbantian and Schotter (1997) use the same approach regarding the number of periods and whether subjects should know the terminal period. At least since Selten’s (1978) chain-store paradox it is well-known that in a finitely repeated Prisoners’ Dilemma the optimal strategy for rational actors in each round is not to cooperate when they assume that the other actors are also rational. Kreps et al. (1982) show that this last assumption is crucial, and that for a rational subject cooperating can be part of an optimal strategy when he or she thinks others might not be rational. Since we are more interested in comparing behavior across treatments than within a treatment, the matter of a finitely versus an infinitely repeated game is only of minor importance given we maintain consistency across treatments.

5 A copy of the instructions can be obtained from the authors on request.
The first of the three treatments was the baseline treatment – the standard Prisoners’ Dilemma (PD) – in which 24 subjects were divided into twelve dyads and the subjects in each dyad played against each other for 25 periods.

In each session of the second treatment (labeled “Direct Democracy”), two groups with five subjects each played the PD with one another. Each member of each group placed an anonymous vote (without preliminary communication) on whether the group should play row 1 or row 2 in Table 2. The group decision was determined via majority rule and every member in the group received the same payoff, independent of his/her vote. Each session included 25 periods of play.

The third treatment (labeled “Representative Democracy”) also is characterized by two groups playing the PD with one another for 25 periods. In this treatment, however, a leader made the decision for his/her group, while the group members voted only on which of two candidates they preferred to be the leader. Each session included two groups with seven subjects – five voters and two candidates (Blue and Yellow in one group, Red and Green in the other). The candidates desired to be the leader not only for decision-making purposes, but also for the 10 tokens they received as payment for being the leader. In every period (except for the first one), a leader chose row 1 or 2, and also decided whether to give a portion of the ten tokens (that he/she was paid) to the electorate. Simultaneously, the other candidate of the same group announced what he/she would have done if he/she had been the leader. After each period, the voters learned their payoff for the period, which was the sum of whatever they received from the PD between the groups (via Table 2) and the tokens transferred to them from the leader. Moreover, they could see the non-binding announcement of the opposition candidate for the same period. The electorate then voted on whether to maintain the incumbent as the leader or to replace him/her with the opposition candidate for the next period. The exception was period 0, in which all candidates announced their intentions if they were selected as the leader for period 1. These announcements were non-binding for all subsequent periods.

4. Results

As previously mentioned, the typical finding in an individual Prisoner’s Dilemma experiment is that subjects tend to cooperate more in early stages of the experiment when compared to latter trials, particularly

---

6 Giving out tokens can be seen as a domestic issue in addition to the international issue. Since we focus on the international side, we do not spell out the domestic one. The reader can think of a public good that a candidate can provide (or promise) which makes him or her worse off but each of the votes better off.

7 Note the distinction between our experiments and Williams’ (1994); while these announcements are costly to observe for the voters in Williams’ set-up, they are free in ours.
when the end period is common knowledge (see, for example, Figure 2 of Andreoni and Miller, 1993). To put our results into perspective, Figures 1 and 2 and Table 3 reveal that this is also true in our baseline treatment, and to an extent in the DD-treatment:

(Figures 1 here)

In the first five periods of the PD-treatment 15% of all outcomes are cooperative, after which cooperative outcomes decline to 5% in periods 6-10. Period 10, however, is the last period in which a dyad cooperates. Similarly, in the DD-treatment, cooperation is most evident in the first five periods (25% of all outcomes are cooperative). Periods 6-21 are characterized by sporadic cooperation, but cooperation is absent in the last four periods:

(Figure 2 here)

(Table 3 here)

A casual comparison of outcomes does not suggest that subjects cooperate more in a direct democracy compared to when they act on their own behalf. Examining the individual actions, however, yields a picture that indicates subjects are more willing to cooperate when they are part of a group. Figure 3 shows the percentage of subjects playing row 2 (cooperation) in the PD. In early rounds cooperation rates are above 50%, but this cooperation rate subsequently decreases until the last few periods, in which subjects completely refrain from cooperative behavior.

(Figure 3 here)

A different outcome is evident in the direct democracy treatments: prior to period 21 a large fraction of subjects vote consistently in favor of cooperation. Only in the last four periods does this strategy dissipate, as the endpoint phenomenon appears to affect behavior, although cooperation does not disappear altogether as some subjects still choose to cooperate.

(Figure 4 here)

8 Sally (1995) finds in his meta-analysis of social dilemmas that the knowledge about the number of periods does not have a significant impact on the rate of cooperation while the number of iterations does – the more periods an experiment has the lower
As such, the regime of direct democracy provides a bit of anecdotal evidence in favor of team or group models – even though the DD-experiment is not a two-level game in that there is no decision-maker on level 1, we see a first indication of the potential of two-level games to provide realistic intuition into models that allow group interactions.

Even more striking is the subject and group behavior in the Representative Democracy treatment. In three of the four sessions we find cooperative outcomes in the intergroup PD in unlikely places (see Figure 5). In one session, cooperation starts in period 11 and continues through period 15; in the second session, cooperation occurs in periods 4, 5, 6, 15, 20 and 22; and in the third session, groups do not cooperate before period 18 but then cooperate until the last period. Only in the fourth session do leaders fail to cooperate in a single period.

(Figure 5 here)

A main catalyst for this observed behavior represents the basic premise underlying the theory of two-level games: because of their obligation to consider domestic interests, governments cannot behave analogously to unitary actors. International issues, however, are only one arena where governments make important decisions. Making choices on several issues simultaneously, politicians have the opportunity to compensate potential losers on one issue with better performance on another. In our simple experiments, the second issue is domestic in nature – determining the optimal token transfer to the electorate. Since a candidate who is not in office can make non-binding announcements, one non-unique Nash equilibrium is for both candidates in a group to give all ten tokens to their voters, or to announce that they would do so (except for the last period, when the optimal token choice for a leader is to keep all tokens since there are no further elections). Figure 6 shows that candidates learned this equilibrium strategy – in each case they were involved in a “race-to-the-bottom” (or “race-to-the-top” from the voters’ perspectives), and gave an increasingly larger portion of their 10 tokens as time passed. Not surprisingly, the opposition candidates typically announced a large transfer if they were voted into office.

(Figure 6 here)

Our results can also be compared to the aforementioned findings of Niskanen (1997) and Kollman et al. (1997). Although Niskanen’s results are not directly comparable to ours due to fundamental differences in the rate of cooperation.
the data generating process, our results are consistent with Niskanen’s findings in one important respect – outcomes in the direct democracy are superior to outcomes in a setting with autocratic subjects. Yet, an interesting contrast is that we find that elections do not hinder the potential for superior outcomes in the representative democracy experiments. Furthermore, concerning multiple jurisdictions, our results support Kollman et al.’s (1997) general finding – representative democracy led to higher utility levels for voters than direct democracy. Overall we observe that combining the two-levels of decision-making stimulates subjects to behave differently from observed patterns of behavior in typical one-level games.

5. Discussion and Conclusion

Most theoretical work on international environmental cooperation uses the paradigm of a country as a unitary rational actor. Despite the merits of such models, it is important to recognize that they fail to account for other interactions inherent in representative decision-making. In this experimental study we used three different experimental treatments – a standard two-player Prisoner’s Dilemma, a direct democracy, and a representative democracy – to compare individual and collective behavior when subjects face different group settings. We find that the propensity to cooperate increases when subjects move from individual games to group games. We also observe an interesting nuance in the representative democracy regime – cooperation is robust to latter periods in the experiment. One potential conjecture to explain this result is that candidates use a second political issue (transfer payments) to compensate the electorate for potential losses in the Prisoner’s Dilemma. Combining these findings with recent results in the social psychology literature suggests that there is a behavioral component inherent in games that group subjects, and this component is potentially incorrectly ignored in traditional game-theoretic models of international interactions.

More work needs to be done in this area. Our study only represents a beginning to a rigorous research agenda. Several extensions are readily apparent:

- allow for more than two options. Often, countries do not only have the choice between cooperation and non-cooperation but can choose different levels of cooperation. The question arises whether domestic constraints would lead to a higher or lower level.
- allow for more than two groups and/or more than one international topic. Folmer et al. (1993) and Susskind (1994) point out that careful crafting of international environmental negotiations by linking them with one other (seemingly unrelated issues like trade, debt or R&D) can induce more successful outcomes. One shortcoming of interconnected games (Folmer et al.) or linked negotiations (Susskind),

---

9 For comparison of the merits and problems of different models in the study of international environmental politics see Skjærseth (1995).
however, is that they might link domestic groups to an issue in which these groups were previously uninterested.

- allow for pre-game chat between the players. Although cheap talk represents an emerging literature, it would be interesting to model such behavior in a two-level game where cheap talk occurs at different levels of the game.

- allow heterogeneous voters and incomplete information about the preferences of voters. Difficulties of politicians in two-level games are a) they face different domestic groups with different preferences, and b) they fail to recognize the strengths of each group, and the preferred position of the median voter.

- allow for periods without elections. Given that politicians behave differently at the beginning of a term when compared to years immediately preceding election time, modeling temporal behavior patterns has consequences on international outcomes since a leader may be more willing to make concessions when elections are further away.

- allow collective and individual interests to differ. Perhaps the most obvious extension is the use of heterogeneous payoff structures.
References


Table 1: Payoff table for cooperation (row 2) versus non-cooperation (row 1) in the abatement game.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.5</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>2.5</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2: Payoff table as used in the experiment.

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>15</td>
<td>5</td>
</tr>
<tr>
<td>2</td>
<td>35</td>
<td>25</td>
</tr>
</tbody>
</table>

Figure 1a: Number of cooperative outcomes in the baseline treatment.
Figure 1b: Number of cooperative outcomes in the Direct Democracy treatment

Figure 2: Percentage of cooperative outcomes in PD, DD, and RD

Table 3: Comparison of cooperative and Nash-outcomes in the three treatments
Figure 3: Percentage of playing Cooperation (row 2) in the baseline treatment

Figure 4: Percentage of votes for cooperation in the Direct Democracy treatment

Figure 5: Number of cooperative outcomes in the Representative Democracy treatment
Figure 6: Average numbers of tokens given by the leader or announced by the opposition