DISCUSSION PAPER SERIES

IZA

IZA DP No. 5189

Group Membership, Competition, and Altruistic versus Antisocial Punishment: Evidence from Randomly Assigned Army Groups

Lorenz Goette
David Huffman
Stephan Meier
Matthias Sutter

September 2010

Forschungsinstitut
zur Zukunft der Arbeit
Institute for the Study
of Labor

Group Membership, Competition, and Altruistic versus Antisocial Punishment: Evidence from Randomly Assigned Army Groups

Lorenz Goette
University of Lausanne and IZA

David Huffman
Swarthmore College and IZA

Stephan Meier
Columbia University and IZA

Matthias Sutter
University of Innsbruck,
University of Gothenburg and IZA

Discussion Paper No. 5189
September 2010

IZA

P.O. Box 7240
53072 Bonn
Germany

Phone: +49-228-3894-0
Fax: +49-228-3894-180
E-mail: iza@iza.org

IZA Discussion Paper No. 5189
September 2010

ABSTRACT

Group Membership, Competition, and Altruistic versus Antisocial Punishment: Evidence from Randomly Assigned Army Groups*

We investigate how group boundaries, and the economic environment surrounding groups, affect altruistic cooperation and punishment behavior. Our study uses experiments conducted with 525 officers in the Swiss Army, and exploits random assignment to platoons. We find that, without competition between groups, individuals are more prone to cooperate altruistically in a prisoner's dilemma game with in-group as opposed to out-group members. They also use a costly punishment option to selectively harm those who defect, encouraging a norm of cooperation towards the group. Adding competition between groups causes even stronger in-group cooperation, but also a qualitative change in punishment: punishment becomes antisocial, harming cooperative and defecting out-group members alike. These findings support recent evolutionary models and have important organizational implications.

JEL Classification: C72, C91, C93

Keywords: group membership, competition, punishment, army, experiment

Corresponding author:

Matthias Sutter
Department of Public Finance
University of Innsbruck
Universitaetsstrasse 15
A-6020 Innsbruck
Austria
E-mail: matthias.sutter@uibk.ac.at

* We thank seminar and conference participants at Harvard University, George Mason University, UC Berkeley, UC Santa Barbara, University of Pittsburgh, Temple University, Tel Aviv University, Wharton School, Tilburg University, ESA Tucson for very helpful comments.

1 Introduction

Many beneficial exchanges require humans to cooperate and trust each other, even though narrow self-interest may tempt them to act selfishly. Evidence suggests, however, that while altruistic cooperation is measurable it is often not strong enough on its own to sustain high levels of cooperation (e.g., Fehr and Schmidt, 1999). Rather, cooperation is often sustained by a willingness of subjects to expend resources to impose harm on others who act selfishly; this altruistic, and costly, informal sanctioning can deter defection (Güth et al., 1982; Roth, 1995; Fehr and Gachter, 2000; Falk et al., 2005). The pervasive tendency for formal contracts to be incomplete (Williamson et al., 1975; Ostrom, 1990; Hechter and Opp, 2001; Knez and Simester, 2001; MacLeod, 2007) underlines the many potential opportunities for informal sanctions to improve efficiency within societies and organizations.

However, the idea that punishment helps enforce a norm of cooperation that leads to efficiency gains, and is therefore a pro-social motivation, is hotly contended (Dreber et al., 2008; Egas and Riedl, 2008; Herrmann et al., 2008; Houser et al., 2008; Abbink et al., 2010). Herrmann et al. (2008) show that in many societies punishment does not punish solely defections from the supposed norm of cooperation, but rather takes the form of antisocial punishment. The latter form of punishment denotes the sanctioning of people who behave cooperatively. Such antisocial punishment does not enhance efficiency, but instead wastes resources of both the punisher and the punished. Rather than sustaining cooperation, antisocial punishment impedes cooperation. The fundamental question addressed in this paper is under which conditions the beneficial or harmful form of punishment manifests? In other words, we examine the determinants of prosocial, altruistic punishment versus antisocial punishment.[1]

[1] We follow the terminology in the literature, and refer to costly punishment of defectors as "altruistic punishment" (Fehr and Gachter, 2000; Herrmann et al., 2008).

The central premise of this paper is that the interaction of social group membership, with the economic conditions of inter-group relationships (mutual independence or competition), is the key to understanding in which situations punishment may be a prosocial act to foster cooperation and when it might turn into efficiency-reducing antisocial punishment. We run experiments with 525 officers in the Swiss Army. For the duration of its officer training program, the Army randomly forms platoons of the officer candidates, thus providing us with a strong, yet exogenous manipulation of group membership. We can show that when there is no competition between groups, individuals cooperate more in a prisoner's dilemma game with in-group members than with out-group members. They also use a punishment option to enforce cooperation norms, especially towards their group. Importantly, we do not find that individuals punish others more simply because they are members of another group. Hence, group boundaries *per se* create a punishment mechanism that favors a norm of cooperation within one's group.

In a second treatment, we add competition between groups to the experiment. In this treatment, the tendency to cooperate with in-group members is even more pronounced. As we explain in section 2, we add the competition between groups such that it leaves the monetary payoffs from punishment unchanged. Thus, any change in punishment behavior must be due solely to the creation of a more competitive atmosphere, not because of a change in which punishment can affect outcomes. We find a qualitative, and dramatic change in punishment patterns: We find strong out-group hostility, in the sense that out-group members are punished harder, and punishment is applied to cooperative and defecting out-group members alike. Thus, punishment here takes on the qualitative pattern of antisocial punishment directed towards the out-group. Thus, our results show that competition triggers a non-strategic desire to harm the out-group, even though this is personally costly. These findings are in line with recent evolutionary models, in which altruism survives partly because, in times of scarcity, it can make people willing to engage in personally costly attacks on competing groups (Choi and Bowles, 2007; Bowles, 2009).

Recently, social group membership effects have received a lot of attention in the economics literature. One the one hand, it has been argued that individuals have social preferences that favor the in-group. So, membership in various groups may affect the willingness of members of an organization to engage in prosocial behavior like altruistic cooperation, or norm enforcement, which enhances efficiency but involves no personal material reward (see, e.g., Akerlof and Kranton, 2000, 2005). On the other hand, it has been conjectured that group membership could have a dark side as well (e.g., Durlauf, 1999) by leading to strong out-group hostility (for a survey, see Hewstone et al., 2002).[2] Sadly, it is obvious that many political conflicts and wars are along (ethnic) group lines (see, e.g., Blattman and Miguel, 2010) and it is a strong empirical regularity that ethnic diversity, i.e. intense interaction of different groups, decreases cooperation (e.g., Alesina et al., 1999; Habyarimana et al., 2007). Interestingly, though, one can think of many examples where group boundaries do not necessarily lead to conflict. Therefore, there has to be an additional factor on top of group membership *per se* that triggers out-group hostility.

Of course, it has already been shown experimentally that competition between groups has behavioral implications for in-group and between-group interactions, leading to more cooperation within groups (Bornstein and Ben-Yossef, 1994; Bornstein et al., 2002; Augenblick and Cunha, 2009). Similar to earlier studies, we also find that competition increases cooperation within groups, without generating a taste for harming the out-group (Halevy et al., 2008). Our results are similar, in that we find hostility is not expressed through cooperation decisions.

We differ from the earlier literature by examining how group boundaries and competition

[2]Note that, conceptually, in-group favoritism need not be the flip-side of out-group hostility. Individuals might favor the in-group, without going out of their way to harm the out-group (see, e.g., Bahry et al. (2005) for evidence of inter-ethnic trust in Russia).

affect the motives behind punishment, moving it closer to the experimental study of conflicts. Research studies using field data have often conjectured that competition explains when groups are in conflict and when they are not. For example, the extent of competition for political power has been argued to explain why the Chewas and the Tubukas are enemies in Malawi, but are friends right across the border in Zambia (Posner, 2004), and conflict between natives and immigrants has been linked to the extent of competition in the job market (Esses et al., 1998). The main challenge for this type of study, however, is the presence of many factors that confound a clean identification of group effects *per se*, and of the effect of the economic environment. For example, groups typically differ according to many characteristics, and these differences could drive behavior rather than group membership. The extent of competition is also often not randomly assigned, raising the concern that more hostile types self-select into competitive situations. Also, behavioral measures have typically not allowed disentangling strategic motives from non-strategic motives that lead to conflict between groups.[3]

Our approach of using Army officers is useful because it keeps many aspects from the field while still leveraging the advantages of experimental methods. We build, in general, on Goette et al. (2006), where we investigated in-group and out-group effects, but did not examine how cooperation and punishment behavior interact with absence or presence of competition. The use of randomly assigned real groups, i.e., platoons, has significant advantages over most previous approaches. Studies based on real groups or existing friends (e.g., Fershtman and Gneezy, 2001; Bahry et al., 2005; Bernhard et al., 2006; Leider et al., 2009) analyze groups with social content and social ties which is an important aspect of real groups. However, these groups are endogenously formed or differ in other

[3]The seminal study in psychology involved young boys being randomly assigned to different groups at a summer camp, and being observed as they first played competitive games, and then engaged in cooperative activities (Sherif et al., 1961). Our study is different because we use adults, and more importantly because we have controlled choice experiments where anonymity allows disentangling strategic and non-strategic motives.

dimensions than just their group membership (i.e. ethnicity), making inferences about the effects of groups *per se* difficult. A solution to these confounds is to randomly assign individuals to so-called “minimal” groups which are nothing more than a label (e.g., Tajfel et al., 1971; Charness et al., 2007; Chen and Li, 2009; Sutter, 2009). By design, these groups lack any social content and behavior might be very different from real groups, as shown in Goette et al. (2010). Combining the advantages of both group manipulation methods, our groups do have social content as groups in the former method, but are at the same time randomly assigned as groups in the minimal group-paradigm.

Our results fit well with previous evidence that antisocial punishment is especially pronounced in societies with more ”close-knit” social networks, where people may perceive everyone outside their network as a competitor (Herrmann et al., 2008). Importantly, in contrast to earlier studies documenting the patterns of altruistic punishment (Fehr and Gachter, 2000) or antisocial punishment (Herrmann et al., 2008; Abbink et al., 2010), we have been able to trigger each by an experimental treatment, thus integrating the seemingly contradictory, earlier findings on the altruistic or antisocial form of punishment in a unified framework. The latter is a main contribution of this paper.

Our results also have important economic implications, in terms of understanding the role of group boundaries, and inter-group competition, within organizations. While other studies have examined the impact of social ties on behavior within organizations (e.g., Bandiera et al., 2010), our study uses assignment to groups as an exogenous manipulation of social ties, and identifies a potential trade-off for firms of fostering competition between social groups: A more competitive scheme may yield more cooperation within groups, but more hostility between groups. Think of a company that lets two teams develop ideas for an advertisement campaign for a new product. A competitive incentive scheme could give a large bonus to the team that comes up with the more convincing campaign, perhaps even firing the other one, while a non-competitive scheme might simply pay the baseline salary whichever team produces the more promising campaign. The competitive scheme might increase efforts and cohesion within teams while limiting the opportunity to exchange ideas across teams, although such an exchange could lead to an even better campaign. Even more so, it might lead to destructive acts of withholding important information for the rival team, although providing it would benefit the company as a whole. Our results show that such a tradeoff can be tipped in favor of one or the other direction by changing the economic environment in which teams act. Thus, there is a delicate balance of cooperation and punishment within and between groups.

The outline of the paper is as follows. In section 2 we introduce the experimental design. Section 3 presents behavioral hypotheses, and section 4 reports the experimental results. Section 5 concludes the paper.

2 Experimental Design

We design experiments in the Swiss Army, which allows us to exploit the random assignment of individuals into platoons as our group manipulation. We then use two experiments and two treatments in each experiment to investigate the effect of competition between groups on cooperation and particularly on norm enforcement.

2.1 Subject Pool and Random Group Assignment

All Swiss males are required to perform at least 300 days of military service, beginning with twenty-one weeks of basic training. In week seven, about one fourth are selected to go through ten weeks of officer-candidate training. Of these, one fourth are promoted to officers and continue on to the Joint Officer Training Program (JOTP).[4] Whereas

[4]The Swiss Army is organized as a reserve system and also officers - after the training - serve only a couple of days per year in the Army.

officer-candidate training is specific for each branch of service, and occurs in separate locations, JOTP brings new officers from all branches of service together, to the same location, for four weeks. Officers are randomly assigned to a platoon at the beginning of JOTP, and spend virtually all time during the day with their platoon. Training involves mainly coursework on principles of security, combat in large military units, logistics, and leadership. At the end of JOTP, the platoons are dissolved and officers are once again sent to separate locations, for further, advanced training specific to each branch of service.

We use the random assignment of candidates to platoons in JOTP as our manipulation of social groups. Each platoon is identified by a different number. Assignment to platoons is random, and stratified according to the different branches of service. The Army assigns platoons orthogonally to any previous social ties among officers with the aim of promoting exchanges of perspectives among different individuals and branches of service.

The assignment mechanism is ideal, in several ways, for investigating the impact of group membership on behavior. First, trainees know that platoon composition is designed to be identical and that nobody could choose which platoon to join. Indeed, statistical tests reveal no significant differences in platoon composition, by branch of service, education, or age. Second, there is no competition between the groups (or trainees) for evaluations or other resources. Relative performance evaluations were completed previously, in candidate training. Thus, there is no function of the group assignment, other than to affect the circle of individuals with whom an officer interacts most frequently. Third, social interactions within a platoon are intense. Platoon members spend the whole workday with their group, for the three weeks leading up to our experiments. Despite the fact that platoons are assigned orthogonally to previous social ties, social interactions and ties also arise endogenously within platoons in after-work time: In a questionnaire, officers in our study report to a question on *"How often do you spend off-duty time with members of a) your own platoon or b) the other platoons?"* that they spend significantly more time off-duty with members of their own platoon. This is remarkable in itself, given that 79.8 percent of the trainees know people in other platoons, mostly from earlier stages of their training. Yet, as illustrated in Table 1, they choose to spend most of what little off-duty time they have with members of their newly assigned group. Thus, platoon assignment provides a strong group manipulation.

[Table 1 about here.]

By using randomly assigned real groups we do not have to rely on arbitrary, minimal groups that lack social ties as a key component of groups. But we also do not have to rely on endogenously formed groups (as in, e.g., Leider et al., 2009) or on groups that differ in other dimensions than just membership to different groups, e.g. nationality or ethnicity (as in, e.g., Bernhard et al., 2006; Habyarimana et al., 2007; Bahry et al., 2005). This allows us to make inference about the causal effect of real groups on behavior.

2.2 Experiments and Group Conditions

In the third week of the four-week training period, we conducted two experiments with the officer candidates to see how random group assignment and random introduction of competition between groups affect behavior. In this subsection we present the two types of experiments and in the next subsection we introduce the two treatments with which we vary the economic environment as being competitive or non-competitive.

Experiment 1: Cooperation. The game was a simultaneous prisoners' dilemma (PD). The players, labeled A1 and A2, were each endowed with 20 points worth real money (4 points = 1 CHF). They simultaneously decided whether to keep the points or pass all of them to the other player. Passed points were doubled. Thus, if both players passed their points (cooperation), they each got 40 points. However, a selfish player could always do better by keeping the points (defecting), regardless of the other player's decision: Defecting when the other defected would yield 20, whereas cooperating would sacrifice the endowment and yield nothing in return; defecting when the other cooperated would yield 60, the maximum possible individual payoff in the game (while leaving the cooperator with 0). Cooperation thus entails lowering one's own payoff, and improving the payoff of the other player, and is an indicator of non-selfish motives. We use the game as our workhorse for studying how group boundaries, and economic environment, affect non-selfish motives for cooperation.

Experiment 1 involved two conditions in a between-subject design. In all conditions, a subject never learned the individual identity of their partner. In the *in-group* condition, subjects interacted anonymously, except for being informed that the other player was a member of their own platoon. The *out-group* condition was the same, except subjects were informed that the other player was a member of another platoon. Group affiliation was clearly marked on the decision sheets. These conditions allow us to examine how group assignment affects cooperation. For a selfish individual, the group affiliation will not change the prediction that he will always defect.

We also elicited individual's beliefs about in-group and out-group cooperation. Independent of the condition they were in, we asked participants to state both their belief for in- and out-group cooperation. We asked them to predict the percentage of the in- and out-group that would send all of the points (cooperate). They were given an incentive to make their best guess: they knew that their prediction would be compared to the percentage actually observed. If the deviation was less than 10 percentage points, then they would get one extra point.

At the very end of the experimental sessions, we conducted a short survey in which we asked participants whether they agreed or disagreed with three statements about trust:

1) *"In general, people can be trusted."*, 2) *"Nowadays, you can't rely on anybody."*, and 3) *"Dealing with strangers, it is better to be cautious before trusting them."*. Participants answered on a 4-point scale (1 "Agree Strongly", 2 "Agree Slightly", 3 "Disagree Slightly", and 4 "Disagree Strongly"). We created an individual variable, $Trust$, by adding the answers to the three questions and assigning a 1 for the least amount of trust and 4 the highest amount of trust per question (answers to question 1 are reversed coded). This is used to help capture individual differences in beliefs about trustworthiness in our statistical analysis.

Experiment 2: Punishment. In Experiment 2, two players A1 and A2 played a PD as in Experiment 1, but we added two additional players, B1 and B2. Each B-player was endowed with 70 points. B1 could assign up to 10 deduction points to A1, and B2 could do the same to A2. Each deduction point subtracted three points from the A-player, and cost the B-player one point of his endowment. The B-players could condition their choices on the actions of A1 and A2. Thus Experiment 2 incorporated the possibility of third-party punishment (Fehr and Fischbacher, 2004), and is suited for examining determinants of whether punishment takes the form of norm enforcement (selectively punishing defection) or antisocial punishment (punishing both cooperation and defection). As punishment is costly, a selfish B-player would never punish.

To examine the impact of group membership on norm enforcement, we varied the composition of players in each game in a between-subject design. For the remainder of the paper, we refer to the group composition in Experiment 2 from B1's perspective. Thus, A1 always refers to the player that the B-player can punish, while we refer to the other A-player as A2. The four different group compositions we implemented are shown in Figure 1.

[Figure 1 about here.]

Varying the group membership of A1 (while keeping constant the group membership of A2) allows us to investigate how the group identity of the person being punished (A1) matters. We also study how punishment varies with the group affiliation of A2, the person affected by A1's actions. Appendix C provides a translation of the instructions for one group composition in the *Neutral Group Environment* treatment.

2.3 Economic Environment Treatments

We used two treatments to analyze the effect of the economic environment on cooperation and punishment behavior within and between groups.

Neutral Group Environment (NG): In this treatment, we used the randomly assigned groups as our group manipulation and varied the group composition as described above. There was no economic competition between the platoons. Results from this treatment were previously presented in Goette et al. (2006).

Competitive Group Environment (CG): We added competition to the 'Neutral Group Environment' treatment by offering a bonus to the platoon that got the highest payoff in the PD stage. The bonus was 20 points for each member if the platoon got the highest average payoff in the PD. In case of a tie between two platoons, the winner was randomly determined. Because the bonus was based on average payoffs for pairs playing the PD, and cooperation maximized payoffs for the pair, cooperation facilitated winning the bonus for the platoon. Importantly, however, the bonus did not change the incentives for a selfish individual: the best strategy for a selfish A-player was still to defect (for the intuition and a formal test, see section 3 on the behavioral hypothesis). Furthermore, in Experiment 2 the bonus was calculated based on the A-player average payoffs *before* deducting any punishment points imposed by the B-players. B-players (and A-players) knew this. Thus the bonus was irrelevant for the choices of the B-players, regardless of whether they were selfish or altruistic. The rules of the game were made clear in the instructions, and we only began the experiment after control questions verified that all participants understood them.

2.4 Experimental Procedures

The experiment was conducted with paper-and-pencil in a large auditorium and lasted 45 minutes. Subjects did not know of the experiment in advance.

Special care was taken to ensure anonymity. First, subjects were never told the identity of their partner(s). Second, they knew that payoffs would be mailed to home addresses ten days after the experiment, so that all participants would only learn the outcome of the experiment *after* JOTP was over and they were no longer with their platoon. These conditions ensured that the experiment was truly one-shot, and that defection was the optimal choice for a selfish individual. For example, subjects did not need to fear reprisal after the experiment if they chose to defect. Additionally, our procedure eliminated the possibility that punishment might have a benefit in future interactions as participants only knew about the outcome of the experiment after the groups got dissolved. Points earned were converted into Swiss Francs (one point = 0.25 CHF) and the subjects earned on average CHF 14.4 (approximately $14). There was no show-up fee.

Overall, 525 subjects participated in the experiments: 228 in the 'Neutral Group Environment' treatment and 297 in the 'Competitive Group Environment' treatment. 281 were assigned the role of A-players and participated in Experiment 1. Half were assigned to the in-group treatment, and half to the out-group treatment. In the few cases in which the groups had an uneven number of A-players, we randomly used the action of some A-players twice to calculate payoffs. The payoff of these players was determined by the decisions associated with the first match. After participating in Experiment 1, these same subjects participated as A-players in Experiment 2. This procedure introduces a possible order effect for the A-players, but choices of the A-players in Experiment 2 are not of interest for our purposes as we analyze cooperation of A-players in Experiment 1 and norm enforcement of B-players in Experiment 2. 244 subjects were assigned the role of B-players. They participated only in Experiment 2, and were assigned to one of four conditions (see Figure 1). We elicited B-players' deduction points using the strategy method, i.e., they specified how many points to deduct from their associated A-player for each possible combination of actions by A1 and A2.

3 Behavioral Hypotheses

This section develops behavioral hypotheses on how the competitive environment might affect cooperation and punishment behavior. If individuals do not have (group-specific) prosocial preferences, individuals will always defect in the PD game, since this is a dominant strategy. Similarly for punishment, a selfish individual would never punish another player as punishment is costly and there is no benefit of punishment in this one-shot interaction.

The competitive environment, i.e. the small bonus in the *CG* treatment, does not change the predictions for a selfish player. The intuition is straightforward: Cooperation never leads to an increased payoff, because it costs 20 points, and the bonus is only 20 points. In fact, our rules for tie-breaking in case two groups have the same number of points imply that individuals always must expect to lose money when cooperating, because the bonus is only 10 in expected terms. Thus, adding competition cannot generate an increase in cooperation rates through selfish incentives; an increase in cooperation under competition must reflect an effect working through non-selfish motives then (see Appendix A for a proof). The competitive environment also has zero impact on punishment choices of a selfish B-player, by construction. The rules of the game are such that the competition is determined without taking into account punishment. Thus, punishment can have no influence on the likelihood of winning the bonus. Hence, our null-hypothesis can be summarized as follows:

H_0 : *With selfish players, defection by A-players and no punishment by B-players will be the dominant strategies - both in NG and in CG.*

Of course, past research has shown that people are not only willing to cooperate and to punish (for surveys, see Fehr and Schmidt, 2003; Meier, 2007) but that they have group specific social preferences (for evidence with minimal groups, see, e.g., Chen and Li, 2009). With group-specific social preferences, a competitive environment can change individuals' behavior. It has been shown that inter-group competition increases intra-group cooperation and coordination within "minimal" groups (Bornstein and Ben-Yossef, 1994; Bornstein et al., 2002) and real, self-selected, groups (Augenblick and Cunha, 2009).

Recent evolutionary models provide an explanation how such group-specific social preferences can survive: In general, the idea of (cultural) group selection argues that a pattern of altruistic cooperation, and altruistic punishment of defectors, can emerge within groups. These altruistic behaviors can survive because they enhance group fitness, and make groups composed of altruists more likely to survive environmental shocks (Henrich, 2004; Boyd et al., 2003). Crucially, altruism must be parochial, or preferentially directed towards own group members, otherwise altruistic groups lose their relative fitness advantage.

Antisocial punishment can emerge, however, with the introduction of competition for resources between groups. In this case the seemingly benign trait of altruism can play a surprising role, because enhancing own-group fitness is not the only way to win: damaging competitor groups is also a viable strategy. In addition to being even more cooperative within their group, altruists might become hostile towards other groups, and use anti-social punishment as a way to damage outsiders. This taste for hostility could survive because it reinforces the relative fitness advantage of groups with altruists (Choi and Bowles, 2007; Bowles, 2009). These arguments are summarized in two alternative hypotheses, one for A-players and one for B-players.

H_1 : *A-players with group-specific social preferences cooperate more often with in-group members in CG than in NG.*

H_2 : *B-players with group-specific social preferences punish out-group members more often in CG than in NG.*

4 Results

We present the results in two steps: first we analyze the impact of group membership and the economic environment on cooperative behavior. Second, we show how group boundaries and a competitive environment affect punishment behavior.

4.1 Cooperation and Beliefs About Cooperation

Panel A in Figure 2 shows the fraction of individuals cooperating as a function of the group composition and the treatment. In general, the figure shows that individuals are willing to cooperate in the PD and that they exhibit in-group favoritism. In the *NG* treatment, there is a significant and large increase in cooperation if individuals are paired with someone from their own platoon rather than another platoon. In fact, cooperation rates are 18 percentage points higher for within-group interactions than between-group. Notably, the lower cooperation rates with out-group members need not indicate hostility, but might simply indicate less willingness to deviate from the dominant selfish strategy when paired with an out-group member. In the CG treatment, favoritism towards the in-group is even more extreme, with cooperation rates being 36 percentage points higher in within-group than between-group interactions. Thus, the increase in cooperation rates among in-group members is stronger in the competitive environment. However, our results also show that out-group cooperation is not decreasing in CG (a pattern also found in different contexts by Rand et al. (2009) and Herrmann et al. (2008)). Thus, competition does not lead individuals to express hostility toward the out-group by defecting more often. This could reflect the limited "expressive value" of defection; individuals might not see defection as a way to unambiguously express hostility, given that it also coincides with the dominant selfish strategy. Overall, the findings on cooperation support the view that in-group favoritism does not necessarily entail hostility towards the out-group (see, e.g., Bahry et al., 2005).

Panel B in Figure 2 shows that the results on cooperation behavior are fully reflected in the individuals' beliefs; people report that they expect in-group favoritism in NG, and significantly greater favoritism in CG.

[Figure 2 and Table 2 about here.]

The results in Figure 2 are also confirmed in logit models of the following form

$$coop_i = \alpha + \gamma_0 IG_i + \gamma_1 IG_i \times CG_i + \delta CG_i + x_i\beta + e_i \tag{1}$$

where *coop* is an indicator variable equal to 1 if individual i cooperates, and zero otherwise. IG is an indicator variable equal to 1 if the individual is paired with another individual from his platoon (*in-group*) and zero if the other player is from another platoon. The indicator variable CG_i is equal to 1 for the 'Competitive Group Environment' treatment, and zero otherwise. In some specifications, we also add control variables x, an index of a person's self-reported trust (explained in Section 2.1), to increase the precision of the estimates. For ease of interpretation, we report marginal effects.

Results in column (1) of Table 2 show that there is a significant overall in-group effect of almost 30 percentage points. Columns (2) and (3) separate the effects of group membership in the two treatments. In the NG treatment, cooperation is about 20 percentage points higher if the interaction is in-group ($p = 0.03$ in column (2), and $p = 0.05$ in column (3)). The strength of the in-group effect depends on the economic environment. The interaction term between IG and CG shows that the cooperation differential in in-group interactions is about 20 percentage points larger when there is competition ($p = 0.07$ in column (2), and $p = 0.021$ in column (3), where we include an index of trust questions).

We can also estimate OLS models[5] similar to those above with the dependent variable being bel_{ik}, which is individual i's belief about the percentage of individuals cooperating in the two group configurations, k. Because we use two observations per individual, we adjust the error terms by clustering on individuals for possible correlations in e_{ik} within individuals. The results, displayed in columns (4) to (6) of Table 2, show that there is a strong overall in-group effect in beliefs, of almost the identical magnitude as observed in behavior ($p < 0.01$; column (4)). We then separate the in-group effect in the two environments. Beliefs about cooperation are significantly higher for in-group pairings in NG, about the same magnitude as we find for behavior. There is a significant interaction with the economic environment: The in-group differential is 13 percentage points larger in CG than in NG ($p < 0.01$ in both specifications). All in-group differentials in beliefs are within a standard deviation of the in-group differentials in cooperation, showing that the individuals had well-calibrated beliefs.

In sum, group membership *per se* creates in-group favoritism, i.e. individuals cooperate

[5]Estimating the same specification with tobit models does not change the results. Results are available from the authors upon request.

more with in-group members than with out-group members. This effect is also reflected in people's beliefs. Randomly adding competition between the groups increases the in-group favoritism even though it does not change the predictions for individuals under the assumption of selfish preferences. This indicates that a competitive environment has an impact of group-specific social preferences. Importantly, competition increases in-group cooperation without reducing out-group cooperation. Thus, just looking at cooperation, one would conclude that competition between groups increases social efficiency.

4.2 Punishment

We now turn to the analysis of B-players' punishment behavior. Figure 3 displays the results for punishment in situations in which A2 cooperated. The figure allows us to highlight two distinct motives related to the group membership. By varying the identity of A1, the person who can be punished, we can see if punishment depends on whether A1 was a member of the punisher's own group (dark lines) or another group (grey lines). By varying the identity of A2, the player who is the potential victim of defection, we can examine if punishment of A1 depends on whether the victim of defection was from the punisher's group (solid lines) or some other group (dashed lines). The figure also distinguishes between whether A1 cooperated or defected.

Panel A displays the results for the *NG* treatment. There is a clear pattern of norm enforcement in the data: A1 is punished more strongly for defection than cooperation. Punishment of A1 also depends on the identity of A2. If A1 defects, the solid lines (A2 from the punisher's group) are always above the dashed lines (A2 from another group). Thus, individuals are especially prone to punish defection if the "victim" of defection is from the in-group. These results are consistent with the prediction that punishers engage in altruistic punishment in a way that enforces a norm of cooperation toward members of their own group. They also mirror the in-group favoritism observed for cooperation behavior. It is also evident from the figure that the identity of A2 does not matter if A1 cooperates. This indicates a lack of hostility in this treatment. Hostility would imply stronger punishment of an A1 that belongs to another group, regardless of what A1 does. As can be seen in the graph, there is essentially no difference as a function of A1's group affiliation. In sum, group boundaries *per se* do not create hostility in punishment.

[Figure 3 about here.]

Turning to the competition treatment in panel B, we see that the punishment choices are starkly different. Most importantly, there is now a clear difference in punishment depending on whether A1 belongs to the punisher's own group or not. Grey lines (A1 is from another group) are clearly above the dark lines (A1 is from the punisher's group). Thus, out-group individuals are punished significantly harder than in-group members, and importantly, this is true no matter whether the individual cooperates or defects (grey lines are above dark lines in both cases). Thus, the introduction of competition leads to substantial antisocial punishment or hostility. Furthermore, there is no relationship between the identity of A2 and punishment in CG, so the tendency to preferentially punish defection against the in-group is no longer present.

The two different economic environments (neutral and competitive) generated qualitatively different patterns of punishment, as is evident in the figure. A formal statistical test confirms this impression: We estimate the following OLS regressions:[6]

$$PP_{ik} = \alpha + \gamma_1 I_i(\text{A1 out-group}) + \gamma_2 I_i(\text{A2 in-group}) + e_i \qquad (2)$$

where PP are the punishment points that individual i assigns in case k. We include two

[6]The results are maintained in tobit regressions and can be obtained upon request.

indicator variables to capture the effect of the group composition on i's punishment; I(A2 in-group) is equal to 1 if player A2 is from the same group as B1 and 0 otherwise and I(A1 out-group) is equal to 1 if player A1 is from another group as B1 and 0 otherwise.

We estimate equation (2) separately for the two cases where A1 cooperates and the two cases where A1 defects, and estimate these again separately for NG and CG. The coefficients across columns are compared in the bottom panel using two-sided χ^2-tests (see Appendix B for a formal expression of the tests). Table 3 displays the result for the case in which A2 cooperates. The table shows that in the NG treatment, i.e. in the neutral environment, we find stronger punishment of defection against a member of one's own group, i.e. A2 is an in-group member ($p = 0.05$, column(1)), but no effect of the identity of A1 on punishment ($p = 0.85$, column (1)). Importantly, in the NG environment, there is no effect of the group composition on punishment of cooperation (column (3)). In contrast, in CG, we observe a different pattern in punishment. This can be seen in columns (2) and (4) of Table 3. The identity of A2 is no longer significant. However, A1 gets punished more heavily, whether he cooperates or defects, if he is from a different group than the punisher, i.e. A1 is out-group ($p < 0.01$ in columns (2) and (4)). Hence, there is substantial antisocial punishment.

The comparison of the coefficients across columns, i.e. the neutral vs. the competitive environment, in the lower panel of Table 3 shows that in treatment CG, A1 is punished more heavily than in NG if he is out-group ($p < 0.01$). This is true for whether A1 defects or cooperates. The different punishment pattern conditional on the identity of A2 is not statistically significant. In sum, we clearly reject the hypothesis that the effect of A1's and A2's group affiliation on punishment are the same across the two treatments ($p < 0.01$ for both, defection and cooperation of A1). And the results show substantial antisocial punishment in the CG treatment.

[Table 3 about here.]

The results are also robust (qualitatively the same, but slightly weaker) to adding the cases in which A2 defected. Obviously, in the case in which A1 cooperates and A2 defects, A1 has a payoff of zero and punishment can't reduce his payoff further. So due to censoring there is no punishment in this case – even in *CG*. Nevertheless, individuals exhibit hostility in punishment also in the case when A1 defects and A2 defects. For reasons of succinctness, the detailed results have been relegated to the Appendix (see Figure A1 and Table A1 there).

Figure 3 additionally indicates that norm enforcement might be weaker in *CG* than in *NG*, i.e. that punishment seems to depend less on A1's behavior in *CG* than in *NG*: While in Panel A of Figure 3 punishment is clearly higher when A1 defects, regardless of the group composition, that relationship is almost completely muted in Panel B. In order to examine the differential in punishment between cooperation and defection, we estimate for each treatment the following equations:

$$PP_{ik} = \alpha + \gamma_1 I_i(\text{A1 out-group}) + \gamma_2 I_i(\text{A2 in-group}) + \gamma_3 I_k(\text{A1 defects}) + e_i \qquad (3)$$

$$\begin{aligned} PP_{ik} =& \alpha + \gamma_1 I_i(\text{A1 out-group}) + \gamma_2 I_i(\text{A2 in-group}) \qquad (4) \\ & + \gamma_4 I_k(\text{A1 defects}) \times I_i(\text{A2 out-group}) + \gamma_5 I_k(\text{A1 defects}) \times I_i(\text{A2 in-group}) + e_i \end{aligned}$$

in which *A1 defects* equals 1 if A1 defects and 0 otherwise. Regression 4 adds two interaction terms for *A1 defects* and whether A2 is an out-group member or an in-group member.

The results (displayed in Table 4) show that defection is more strongly punished than cooperation in *NG* ($p < 0.01$, column (1)). This effect is less strong in *CG* as seen in column (2). A formal test comparing the two coefficients, γ_3, shows that the difference is significant ($p < 0.01$) and conditioning of punishment on actions of A1 is much weaker in *CG*, while it is still significant. As can be seen from comparing coefficient γ_4 across columns (3) and (4), this effect of competition in *CG* even prevails when A2 is from the punisher's own group ($p = 0.03$), the case when norm enforcement was strongest in *NG*. Finally, it is noteworthy from column (4) that the norm enforcement pattern of punishment if defection is directed against one's own group has completely disappeared in the competitive environment. The fact that individuals cease to use punishment to enforce cooperation among in-group members in a competitive environment might seem surprising, from the perspective of increasing group fitness. However, our results on cooperation showed that competition leads to very high within-group cooperation rates, even in the absence of punishment threat, mitigating this problem for the group. In summary, competition causes punishment to stop functioning as a tool for norm enforcement, and instead to take the form of antisocial punishment directed towards the out-group, consistent with predictions given group-specific social preferences.

[Table 4 about here.]

5 Conclusions

Punishment is a double-edged sword. While it can have positive effects on the level of cooperation within groups by enforcing a norm to contribute to a group's welfare (e.g., Fehr and Gachter, 2000), it can also be detrimental for society by sanctioning cooperative behavior (e.g., Herrmann et al., 2008). In this paper we have investigated the conditions under which the upside (altruistic) or downside (antisocial) of punishment prevails. Our results have shown that the interaction of group boundaries, and the economic environment surrounding the groups, can generate two starkly different patterns of cooperation and punishment behaviors. In the absence of competition between groups, individuals cooperate more within their group, and use punishment to enforce cooperation norms towards their group. Group boundaries *per se* do therefore not lead to inherent hostility towards others in punishment. However, the punishment patterns change dramatically when we introduce competition between groups. We find strong out-group hostility, in the form of antisocial punishment of the out-group, but also increased in-group cooperation. These findings support evolutionary models predicting hostility and aggression between groups in competitive settings (Choi and Bowles, 2007; Bowles, 2009).

Our results provide insights into the origins of conflict along group lines, and have important economic implications in terms of understanding the role of group boundaries and inter-group competition within organizations. While Bandiera et al. (2010) show the impact of social ties on behavior within an organization, we use random assignment to groups as an exogenous manipulation of social ties and show that the economic environment can elicit substantially different behavior with respect to cooperation and punishment. This implies that organizations might face a tradeoff, for instance, when designing incentive schemes for work-teams. While increasing competition between teams increases cooperation within teams it also changes the nature of informal sanctions. The sanctioning mechanism not only loses its often praised benefits but takes the form of between-group hostility potentially outweighing benefits from increased in-group cooperation.

Our results also provide additional evidence that social preferences are endogenous to the economic environment (Bowles, 1998; Burks et al., 2009). This literature argues that changes in economic environments bring about changes in preferences. Yet, these changes are typically assumed to be slow, e.g., operating through slow-changing norms of cooperation (Herrmann et al., 2008). It is noteworthy that we observe a particularly strong form of endogenous preferences: Our treatments are between-subject manipulations, and still we immediately observe starkly different punishment and cooperation strategies, conditional on the economic environment. Thus, this evidence suggests that different motives of social preferences may be dormant in humans, and triggered by different economic environments.

References

Abbink, K., J. Brandts, B. Herrmann, and H. Orzen (2010): "Intergroup Conflict and Intra-Group Punishment in an Experimental Contest Game," *The American Economic Review*, 100, 420–447.

Akerlof, G. and R. Kranton (2005): "Identity and the Economics of Organizations," *Journal of Economic Perspectives*, 19, 9 – 32.

Akerlof, G. A. and R. E. Kranton (2000): "Economics and identity," *Quarterly Journal of Economics*, 115, 715–753.

Alesina, A., R. Baqir, and W. Easterly (1999): "Public Goods and Ethnic Divisions," *Quarterly Journal of Economics*, 114, 1243–1284.

Augenblick, N. and J. Cunha (2009): "Using competition to elicit cooperation in a political public goods game: A field experiment," *Working Paper*.

Bahry, D., M. Kosolapov, P. Kozyreva, and R. Wilson (2005): "Ethnicity and trust: Evidence from Russia," *American Political Science Review*, 99, 521–532.

Bandiera, O., I. Barankay, and I. Rasul (2010): "Social incentives in the workplace," *Review of Economic Studies*, 77, 417–458.

Bernhard, H., E. Fehr, and U. Fischbacher (2006): "Group Affiliation and Altruistic Norm Enforcement," *American Economic Review*, 96, 217–221.

Blattman, C. and E. Miguel (2010): "Civil War," *Journal of Economic Literature*, 48, 3–57.

Bornstein, G. and M. Ben-Yossef (1994): "Cooperation in inter-group and single-group social dilemmas," *Journal of Experimental Social Psychology*, 30, 52–52.

Bornstein, G., U. Gneezy, and R. Nagel (2002): "The effect of intergroup competition on group coordination: an experimental study," *Games and Economic Behavior*, 41, 1–25.

Bowles, S. (1998): "Endogenous preferences: The cultural consequences of markets and other economic institutions," *Journal of Economic Literature*, 36, 75–111.

——— (2009): "Did Warfare Among Ancestral Hunter-Gatherers Affect the Evolution of Human Social Behaviors?" *Science*, 324, 1293–1298.

Boyd, R., H. Gintis, S. Bowles, and P. J. Richerson (2003): "The evolution of altruistic punishment," *Proceedings of the National Academy of Sciences*, 100, 3531–3535.

Burks, S. V., J. P. Carpenter, and L. Goette (2009): "Performance pay and worker cooperation: Evidence from an artefactual field experiment," *Journal of Economic Behavior and Organization*.

Charness, G., L. Rigotti, and A. Rustichini (2007): "Individual Behavior and Group Membership," *American Economic Review*, 97, 1340–52.

CHEN, Y. AND S. X. LI (2009): "Group identity and social preferences," *American Economic Review*, 99.

CHOI, J.-K. AND S. BOWLES (2007): "The coevolution of parochial altruism and war," *Science*, 318, 636.

DREBER, A., D. G. RAND, D. FUDENBERG, AND M. A. NOWAK (2008): "Winners don't punish," *Nature*, 452, 348–351.

DURLAUF, S. (1999): "The Case 'Against' Social Capital," *Focus*, 20, 1–5.

EGAS, M. AND A. RIEDL (2008): "The economics of altruistic punishment, and the maintenance of cooperation," *Proceedings of the Royal Society*, 275, 871 – 878.

ESSES, V. M., L. M. JACKSON, AND T. L. ARMSTRONG (1998): "Intergroup competition and attitudes toward immigrants and immigration: An instrumental model of group conflict," *Journal of Social Issues*, 54, 699–724.

FALK, A., E. FEHR, AND U. FISCHBACHER (2005): "Driving forces behind informal sanctions," *Econometrica*, 73, 2017–2030.

FEHR, E. AND U. FISCHBACHER (2004): "Third-party punishment and social norms," *Evolution and Human Behavior*, 25, 63–87.

FEHR, E. AND S. GACHTER (2000): "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90, 980–994.

FEHR, E. AND K. SCHMIDT (1999): "A theory of fairness, competition, and cooperation," *Quarterly Journal of Economics*, 114, 817–868.

——— (2003): "Theories of Fairness and Reciprocity - Evidence and Economic Application," in *Advances in Economics and Econometrics - 8th World Congress, Econometric Society Monographs*, ed. by M. Dewatripont, L. P. Hansen, and S. J. Turnovsky, Cambridge: Cambridge University Press, 208–57.

FERSHTMAN, C. AND U. GNEEZY (2001): "Discrimination in a Segmented Society: An Experimental Approach," *Quarterly Journal of Economics*, 116, 351–377.

GOETTE, L., D. HUFFMAN, AND S. MEIER (2006): "The Impact of Group Membership on Cooperation and Norm Enforcement: Evidence using Random Assignment to Real Social Groups," *American Economic Review (Papers & Proceedings)*, 96, 212 – 216.

——— (2010): "The Impact of Social Ties on Group Interactions: Evidence from Minimal Groups and Randomly Assigned Real Groups," *Working Paper*.

GÜTH, W., R. SCHMITTBERGER, AND B. SCHWARZE (1982): "An experimental analysis of ultimatum bargaining," *Journal of Economic Behavior & Organization*, 3, 367–388.

HABYARIMANA, J., M. HUMPHREYS, D. POSNER, AND J. WEINSTEIN (2007): "Why does ethnic diversity undermine public goods provision?" *American Political Science Review*, 101, 709–725.

Halevy, N., G. Bornstein, and L. Sagiv (2008): "'In-Group Love' and 'Out-Group Hate' as Motives for Individual Participation in Intergroup Conflict: A New Game Paradigm," *Psychological Science*, 19, 405–411.

Hechter, M. and K.-D. Opp (2001): *Social norms*, Russell Sage Foundation Publications.

Henrich, J. (2004): "Cultural group selection, coevolutionary processes and large-scale cooperation," *Journal of Economic Behavior and Organization*, 53, 3–35.

Herrmann, B., C. Thoni, and S. Gachter (2008): "Antisocial Punishment Across Societies," *Science*, 319, 1362–1367.

Hewstone, M., M. Rubin, and H. Willis (2002): "Intergroup Bias," *Annual Review of Psychology*, 53, 575 – 604.

Houser, D., E. Xiao, K. McCabe, and V. Smith (2008): "When punishment fails: Research on sanctions, intentions and non-cooperation," *Games and Economic Behavior*, 62, 509–532.

Knez, M. and D. Simester (2001): "Firm-wide incentives and mutual monitoring at Continental Airlines," *Journal of Labor Economics*, 19, 743–772.

Leider, S., M. Möbius, T. Rosenblat, and Q.-A. Do (2009): "Directed altruism and enforced reciprocity in social networks," *Quarterly Journal of Economics*, 124, 1815–1851.

MacLeod, W. B. (2007): "Can Contract Theory Explain Social Preferences?" *American Economic Review*, 97, 187–192.

Meier, S. (2007): "A Survey on Economic Theories and Field Evidence on Pro-Social Behavior," in *Economics and Psychology: A Promising New Cross-Disciplinary Field*, ed. by B. S. Frey and A. Stutzer, Cambridge: MIT Press, 51–88.

Ostrom, E. (1990): *Governing the commons: The evolution of institutions for collective action*, Cambridge: Cambridge University Press.

Posner, D. (2004): "The political salience of cultural difference: Why Chewas and Tumbukas are allies in Zambia and adversaries in Malawi," *American Political Science Review*, 98, 529–545.

Rand, D. G., T. Pfeiffer, A. Dreber, R. W. Sheketoff, N. C. Wernerfelt, and Y. Benkler (2009): "Dynamic remodeling of in-group bias during the 2008 presidential election," *Proceedings of the National Academy of Sciences*, 106, 6187–6191.

Roth, A. E. (1995): "Introduction to Experimental Economics," in *The Handbook of Experimental Economics*, ed. by J. H. Kagel and A. E. Roth, Princeton University Press, chap. 1, 3–109.

Sherif, M., O. J. Harvey, J. White, W. Hood, and C. Sherif (1961): *Intergroup conflict and cooperation: The robbers cave experiment*, Norman, OK: University of Oklahoma Book Exchange.

Sutter, M. (2009): "Individual behavior and group membership: Comment," *American Economic Review*, 99, 2247–57.

Tajfel, H., M. Billig, R. Bundy, and C. Flament (1971): "Social categorization and inter-group behavior," *European Journal of Social Psychology*, 1, 149–177.

Williamson, O. E., M. L. Wachter, and J. E. Harris (1975): "Understanding the employment relation: The analysis of idiosyncratic exchange," *The Bell Journal of Economics*, 6, 250–278.

Tables and Figures

Table 1: Off-Duty Time Spent per Week

	Own Platoon	Other Platoon
Less than once	4.5%	30.4%
Between 1 and 2 times	45.0%	44.0%
Twice or more	50.5%	25.7%
N	489	491

Note: The two distributions are significantly different (Wilcoxon Signed-Rank test, $p < 0.001$).

Table 2: Results for Cooperation Rates and Beliefs about Cooperation

Dependent Variable:	**Cooperation (=1)**			**Beliefs: % cooperating**		
	(1)	(2)	(3)	(4)	(5)	(6)
Ingroup (=1)	0.28***	0.18**	0.17*	0.26***	0.18***	0.18***
	(0.05)	(0.08)	(0.09)	(0.02)	(0.02)	(0.02)
CG (=1)	0.08	0.00	0.01	0.01	-0.06*	-0.06*
	(0.06)	(0.08)	(0.08)	(0.03)	(0.03)	(0.03)
Ingroup×CG		0.18*	0.22**		0.13***	0.13***
		(0.10)	(0.10)		(0.03)	(0.04)
Trust			0.06***			0.03***
			(0.02)			(0.01)
Constant				0.37***	0.41***	0.34***
				(0.02)	(0.02)	(0.04)
(Pseudo)-R^2	0.07	0.08	0.10	0.20	0.21	0.25
# of observations	281	281	267	538	538	515
# of individuals	281	281	267	274	274	262

Notes: In column (1) to (3), marginal effects from logit models. In columns (3) and (4) coefficients from OLS models. The model in columns (4) to (6) uses two observations per individual (if available), therefore standard errors of the estimates in column (4) to (6) are adjusted for clustering on individuals.

Level of significance: $^*: 0.05 \le p < 0.1$, $^{**}: 0.01 \le p < 0.05$, $^{***}: p < 0.01$.

Table 3: Punishment as a Function of Group Membership

	(1)	(2)	(3)	(4)
Behavior of A1:	A1 defects		A1 cooperates	
Environment:	Neutral	Comp	Neutral	Comp
A1 out-group (γ_1)	0.155	3.742***	–0.099	2.898***
	(0.853)	(0.658)	(0.693)	(0.684)
A2 in-group (γ_2)	1.694**	0.868	0.535	–0.544
	(0.840)	(0.676)	(0.697)	(0.693)
Constant	4.487***	1.636***	2.307***	1.988***
	(0.705)	(0.447)	(0.578)	(0.517)
R^2	0.039	0.203	0.005	0.125
# of observations/individuals	111	132	111	132
Tests across equations (environments):				
Test that γ_1 differs	$p < 0.01$		$p < 0.01$	
Test that γ_2 differs	$p = 0.44$		$p = 0.27$	
Test that γ_1 and γ_2 differ	$p < 0.01$		$p < 0.01$	

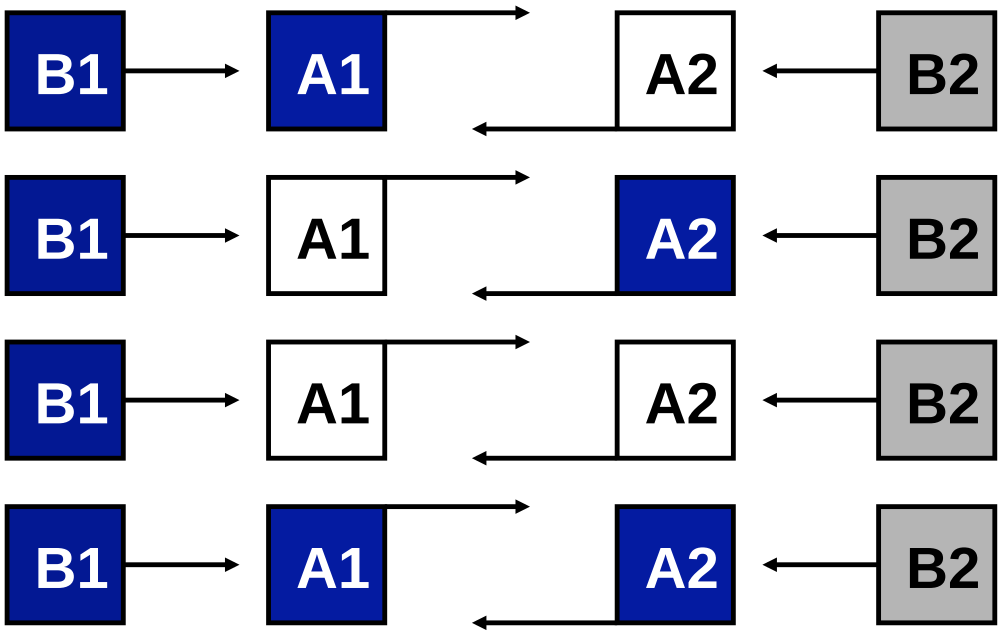
Note: Dependent variable: # of deduction points. OLS estimates for the cases in which A2 cooperates. Robust standard errors in parentheses. p-values in cross-equation tests are all two-sided.

Level of significance: $^*: 0.05 \leq p < 0.1$, $^{**}: 0.01 \leq p < 0.05$, $^{***}: p < 0.01$.

Table 4: Norm Enforcement Across Environments

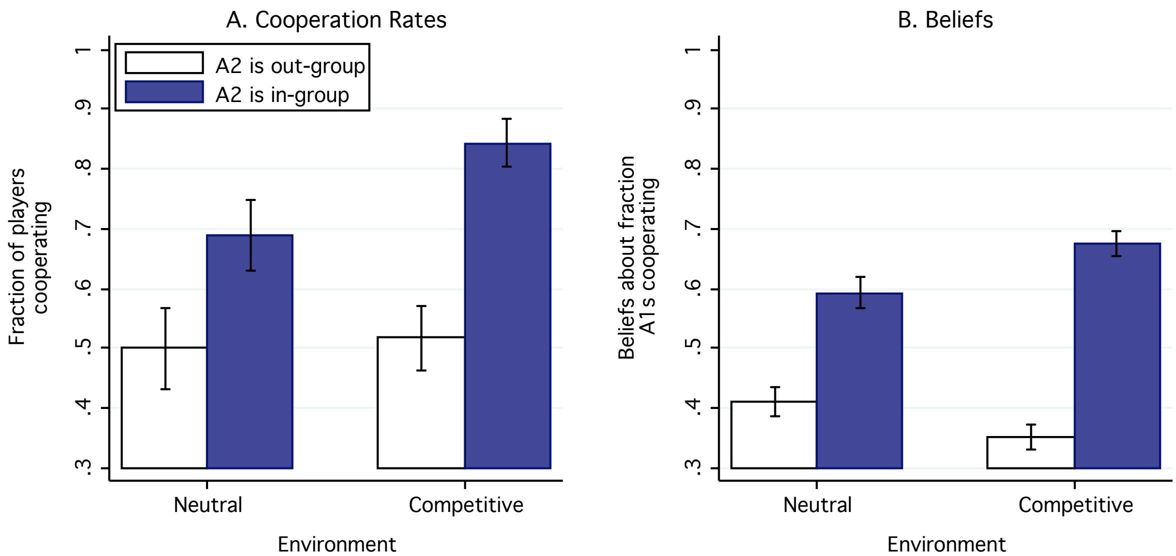
Environment:	Neutral (1)	Comp (2)	Neutral (3)	Comp (4)
A1 out-group (γ_1)	0.028	3.320***	0.028	3.320***
	(0.624)	(0.589)	(0.625)	(0.590)
A2 in-group (γ_2)	1.114*	0.162	0.513	–0.533
	(0.609)	(0.595)	(0.691)	(0.694)
A1 defects (γ_3)	2.838***	0.682**		
	(0.445)	(0.330)		
A1 defects × A2 out-group (γ_4)			3.510***	1.482***
			(0.722)	(0.560)
A1 defects × A2 in-group (γ_5)			2.306***	0.092
			(0.549)	(0.385)
Constant	1.978***	1.471***	2.244***	1.766***
	(0.565)	(0.465)	(0.578)	(0.495)
R^2	0.130	0.161	0.136	0.167
# of observations	222	264	222	264
# of individuals	111	132	111	132
Tests across equations (environments):				
Test that γ_1 differs	$p < 0.01$		$p < 0.01$	
Test that γ_2 differs	$p = 0.26$		$p = 0.29$	
Test that γ_3 differs	$p < 0.01$			
Test that γ_4 differs			$p = 0.03$	
Test that γ_5 differs			$p < 0.01$	
Test that γ_1, γ_2, and γ_3 differ	$p < 0.01$			
Test that γ_1, γ_2, γ_4, and γ_5 differs			$p < 0.01$	

Note: Dependent variable: # of deduction points. OLS estimates. Robust standard errors clustered on the individual in parentheses.
Level of significance: $^{*}: 0.05 \leq p < 0.1$, $^{**}: 0.01 \leq p < 0.05$, $^{***}: p < 0.01$. p-values in cross-equation tests are all two-sided.



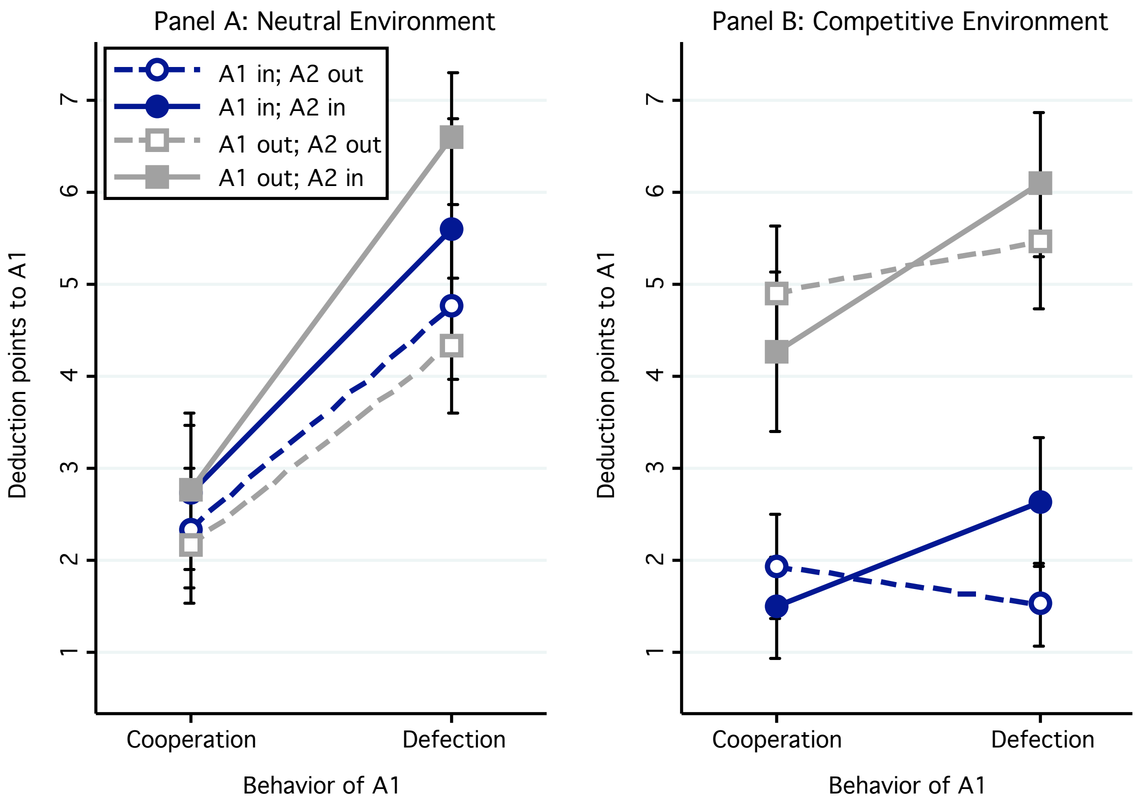
Note: The game allowed B1 to punish A1, and B2 to punish A2, conditional on the actions and A1 and A2 in a simultaneous prisoners' dilemma game. The dark shading indicates the four possible group combinations for B1, A1, and A2, which were implemented as different treatments (players with the same shading are from the same group). The design deliberately did not vary all possible combinations of B1 and B2 group roles, because of number of observations, so the effect of B2 group identity on B1 behavior is not studied. The pattern of B-player (and A-player) group compositions was identical across the Economic Environments, the *NG* and *CG* treatments.

Figure 1: Group Composition in Third-Party Punishment Game



Notes: The bars show standard errors of the mean. White indicates an out-group pairing while dark indicates an in-group pairing. Panel A shows the fraction of A-players passing their endowment to the other player in a simultaneous one-shot PD. Panel B shows beliefs about the fraction of A-players who pass their endowment in the PD.

Figure 2: Cooperation rates and beliefs about cooperation in *NG* and *CT*



Notes: B-players could deduct between 0 and 10 points. Each deduction point costs B-players 1 point and A1-players 3 points. Deduction points were made conditional on whether A1-players cooperated or defected using the strategy method. Error bars show standard errors of the mean.

Figure 3: Punishment in the case A2-player cooperated

Appendix

A Nash Equilibrium in the PD in *Competitive Group Environment*

In this section, we explain why it was optimal for selfish A-players to defect in the Competitive Group Environment. The intuition is straightforward: Cooperating never leads to an increased payoff, because cooperating costs 20 points, and the bonus is only 20 points. In fact, our rules for tie-breaking in case two groups have the same number of points imply that individuals always lose money when cooperating, because the bonus is only 10 in expected terms. Thus, adding competition cannot generate an increase in cooperation rates through selfish incentives; an increase in cooperation under competition must reflet an affect working through non-selfish motives. Below, we formalize the intuition that the Nash equilibrium doesn't change between NG and CG.

We show that the Nash equilibrium in the game involves all (selfish) A-players defecting.

(i) In our experiment, there are within-group and between-group pairings. Obviously, a selfish player never cooperates with a player from another group, since, on top of costing him 20 points, he may also be pivotal in losing the bonus. Therefore, what remains to be considered are within-group pairings. First consider the case of $K = 2$ groups, denoted X and Y. Now pick an arbitrary collection of strategies in which some individuals cooperate in within-group pairings. We ask whether this strategy can be a Nash equilibrium. Two possible cases can arise: Either one of the groups, say group Y, loses, or the two groups tie.

We first show that groups can never tie with some individuals cooperating.

- Pick an arbitrary member of group k who is cooperating. Since the groups are tied, he wins a bonus with probability 0.5. If he defects, his group will lose for sure. However, defecting saves 20 points, while costing only 10 points in expected bonus. Thus, when two groups are tied, cooperating players have an incentive to defect.

We now show that it is impossible to have a Nash equilibrium in which group Y loses for sure.

- If group Y loses, then it cannot be a Nash equilibrium for anyone in group Y to cooperate. Given the others' strategies, members of Y who cooperate can increase their payoff by 20 points if they defect.
- Given this result, it follows that in group X, at most one player will cooperate. If more than one player in group X cooperated, a player could switch to defection while still winning the bonus, holding the other players' strategies constant.
- However, if one player in X cooperates, the tying rule now implies that the player can defect, and save 20 points, but only lose 10 points in expected bonus (since the two groups now tie).

Thus, the only equilibrium for $K = 2$ groups involves both groups tying, and this equilibrium involves all players defecting.

(iii) The above arguments immediately generalize to $K > 2$ groups. The only difference is that the expected bonus in the case of a tie will be even smaller, $20/m$, where $m \leq K$ is the number of groups tying. Thus, the same reasoning applies.

B Comparisons across Equations

The bottom panels of Table 3 to 4 also display cross-equation tests. For single-coefficient tests, we calculate

$$z = \frac{\gamma_j - \tilde{\gamma}_j}{\sqrt{\Sigma_{jj} + \tilde{\Sigma}_{jj}}} \tag{5}$$

where γ_j and $\tilde{\gamma}_j$ are the two coefficients of interest from the two equations, and Σ_{jj} and $\tilde{\Sigma}_{jj}$ are the corresponding main diagonal elements in the covariance matrix (because the two coefficients come from two separate equations, their covariance, by construction, is zero). z has a standard normal distribution under the null of no difference. We report two-sided p-values to be conservative. In the case of coefficient vectors, we calculate the analogous test statistic

$$\chi = (\gamma - \tilde{\gamma})(\Sigma + \tilde{\Sigma})^{-1}(\gamma - \tilde{\gamma})' \tag{6}$$

which has a chi-square distribution with k degrees of freedom, where k is the number of variables in γ.

C Appendix with Experimental Instructions

C.1 Instructions Player A (Translation)

What is this about?

Two subjects participate in this decision situation. They will be called A1 and A2. Both, A1 and A2, will get an endowment of 20 points. Each participants has to decide between two options:

- **Keep**: The participant keeps his 20 points.
- **Transfer**: The participant transfers his 20 points to the other participants. The transfered points will be doubled.

Each participant has to decide whether to Keep or the Transfer without knowing how the other participant decided. So, the following payoffs can result:

		Payoffs in this case:
Case 1:	A1 keeps the points	A1: 20 points
	A2 keeps the points	A2: 20 points
		Payoffs in this case:
Case 2:	A1 transfers 20 points	A1: 0 points
	A2 keeps the points	A2: 60 points
		Payoffs in this case:
Case 3:	A2 keeps the points	A1: 60 points
	A2 transfers 20 points	A2: 0 points
		Payoffs in this case:
Case 4:	A1 transfers 20 points	A1: 40 points
	A2 transfers 20 points	A2: 40 points

How will you decide?

- You will be in the role of A1.
- Your assigned **participant A2 is from another platoon**.

None of the participants will ever find out to whom he was assigned. We guarantee total anonymity. When all the participants reached a decision, we will calculate the points and the resulting monetary payoffs in the following way:

4 points = CHF1

The amount will be delivered to you by mail.

Everything clear?

Before you decide, answer the following questions. The question make sure that all the participants understand the instructions.

If you have questions, please contact the staff.

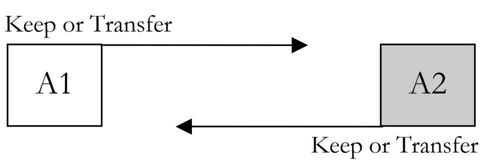
1. A1 and A2 keep their points. Please calculate the resulting points for all participants. State all the steps in getting to the result.
2. A1 and A2 transfer their points. Please calculate the resulting points for all participants. State all the steps in getting to the result.
3. A1 keeps his points and A2 transfer his points. Please calculate the resulting points for all participants. State all the steps in getting to the result.

Please contact the staff when you are done with the questions or if you have questions.

Decision Sheet

- You were assigned the role of **A1**.
- Your assigned **participant A2 is from another platoon**.

In the following figure are the participants from the other platoons shaded.



Please decide which option to pick:

- Keep
- Transfer

Please let the staff know when you decided.

C.2 Instructions Player B (Translation)

What is this about?

Four subjects participate in this decision situation. They will be called A1, A2, B1 and B2. The decision situation will have two steps.

Step 1: A1 and A2 will get an endowment of 20 points. Each participants has to decide between two options:

- **Keep**: The participant keeps his 20 points.
- **Transfer**: The participant transfers his 20 points to the other participants. The transfered points will be doubled.

Each participant has to decide whether to Keep or the Transfer without knowing how the other participant decided. So, the following payoffs can result:

		Payoffs in this case:
Case 1:	A1 keeps the points	A1: 20 points
	A2 keeps the points	A2: 20 points
		Payoffs in this case:
Case 2:	A1 transfers 20 points	A1: 0 points
	A2 keeps the points	A2: 60 points
		Payoffs in this case:
Case 3:	A2 keeps the points	A1: 60 points
	A2 transfers 20 points	A2: 0 points
		Payoffs in this case:
Case 4:	A1 transfers 20 points	A1: 40 points
	A2 transfers 20 points	A2: 40 points

Step 2: B1 and B2 will get an endowment of 70 points each and A1 and A2 will get another 10 points each. In Step 2, B1 and B2 can assign deduction points. B1 can assigned deduction points to A1 and B2 can assign deduction points to A2. B1 and B2 can each assign a maximum of 10 deduction points.

Before explaining how B1 and B2 will make their decisions, we will describe how deduction points will change the payoffs. **Each deduction point will reduce the payoff of B by one point and the payoff of A by three points.** For example, if B1 assigns 3 deduction points, this will reduce A1's payoff by 9 points and B1's payoff by 3 points.

B1 and B2 will decide about the assignment of deduction points for each potential case in Step 1. That is, they will decide about assigning deduction points for the following four potential cases in Step 1:

- **Case 1**: A1 and A2 keep their points.
- **Case 2**: A1 transfers his points and A2 keeps his points.
- **Case 3**: A1 keeps his points and A2 transfers his points.
- **Case 4**: A1 and A2 transfer their points.

This will lead to the following payoffs:

How will you decide?

- You will be in the role of **B1**.
- Your assigned **participant A1 is from another platoon**.

Payoff of A1 =	Payoff from Step 1 + 10 points from Step 2 - 3*Deduction points from B1
Payoff of A2 =	Payoff from Step 1 + 10 points from Step 2 - 3*Deduction points from B2
Payoff of B1 =	Endowment of 70 points - Deduction points to A1
Payoff of B2 =	Endowment of 70 points - Deduction points to A2

- The participant **A2** is **from another platoon**. He got assigned to a participant **B2 from your platoon**.

None of the participants will ever find out to whom he was assigned. We guarantee total anonymity. When all the participants reached a decision, we will calculate the points and the resulting monetary payoffs in the following way:

4 points = CHF1

The amount will be delivered to you by mail.

Everything clear?

Before you decide, answer the following questions. The question make sure that all the participants understand the instructions.

If you have questions, please contact the staff.

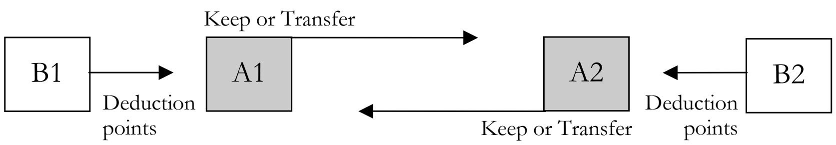
1. In Step 1, A1 and A2 keep their points. In Step 2, neither B1 nor B2 assign any deduction points. Please calculate the resulting points for all participants. State all the steps in getting to the result.
2. In Step 1, A1 and A2 transfer their points. In Step 2, neither B1 nor B2 assign any deduction points. Please calculate the resulting points for all participants. State all the steps in getting to the result.
3. In Step 1, A1 keeps his points and A2 transfers his points. In Step 2, B1 assigns 2 deduction points and B2 assigns 5 deduction points. Please calculate the resulting points for all participants. State all the steps in getting to the result.
4. In Step 1, A1 transfers his points and A2 transfers his points. In Step 2, B1 assigns 1 deduction points and B2 assigns 4 deduction points. Please calculate the resulting points for all participants. State all the steps in getting to the result.

Please contact the staff when you are done with the questions or if you have questions.

Decision Sheet

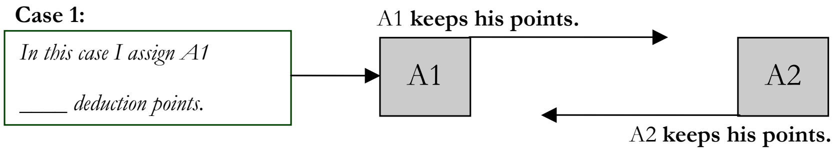
- You will be in the role of **B1**.
- Your assigned **participant A1 is from another platoon**.
- The participant **A2** is **from another platoon**. He got assigned to a participant **B2 from your platoon**.

In the following figure are the participants from the other platoons shaded.

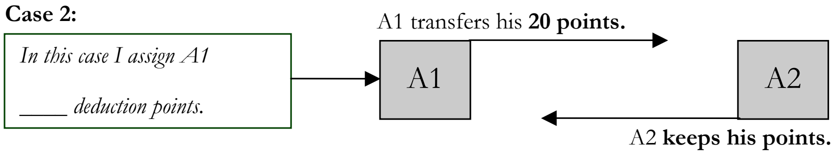


Please decide about the assignment of the deduction points for all possible cases. Only the cases that really happen will determine your payoff and the payoff of the other participants. In each of the cases, you can assign between 0 and 10 deduction points.

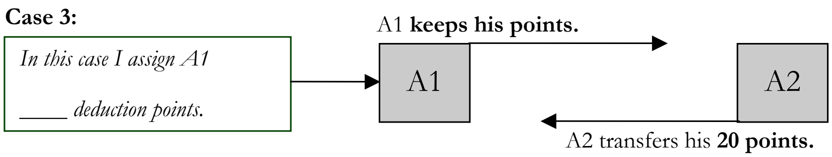
Case 1:



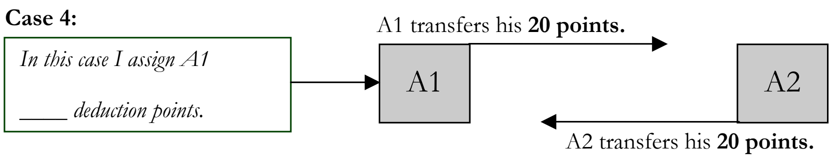
Case 2:



Case 3:

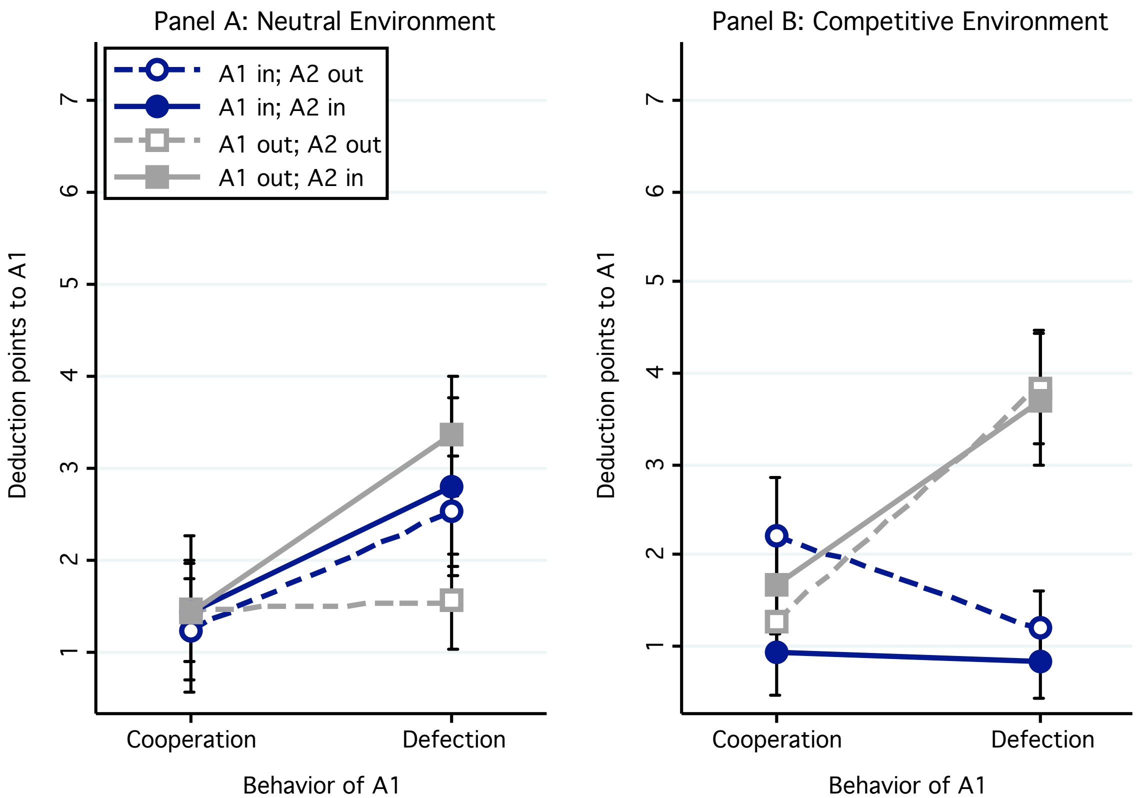


Case 4:



Please let the staff know when you decided.

D Appendix Tables and Figures



Notes: B-players could deduct between 0 and 10 points. Each deduction point costs B-players 1 point and A1-players 3 points. Deduction points were made conditional on whether A1-players cooperated or defected using the strategy method. Error bars show standard errors of the mean.

Figure A1: Punishment in the case A2-player defected

Table A1: Punishment as a Function of Group Membership

	(1)	(2)	(3)	(4)
Behavior of A1:	A1 defects		A1 cooperates	
Environment:	Neutral	Comp	Neutral	Comp
A1 outgroup (γ_1)	–0.103	3.249***	0.018	1.336***
	(0.661)	(0.549)	(0.524)	(0.470)
A2 ingroup (γ_2)	1.445**	0.308	0.297	–0.485
	(0.645)	(0.558)	(0.534)	(0.469)
A2 defects (=1)	–2.774***	–1.477***	–1.084***	–1.705***
	(0.387)	(0.269)	(0.376)	(0.401)
Constant	4.746***	2.127***	2.345***	2.768***
	(0.622)	(0.420)	(0.460)	(0.438)
R^2	0.141	0.205	0.028	0.091
# of observations	223	264	223	264
# of individual	112	132	112	132
Tests across equations (environments):				
Test that γ_1 differs	$p < 0.01$		$p = 0.06$	
Test that γ_2 differs	$p = 0.18$		$p = 0.27$	
Test that γ_1 and γ_2 differ	$p < 0.001$		$p = 0.10$	

Note: Dependent variable: # of deduction points. OLS estimates. Robust standard errors clustered on the individual in parentheses. p-values in cross-equation tests are all two-sided.
Level of significance: $^*: 0.05 \leq p < 0.1$, $^{**}: 0.01 \leq p < 0.05$, $^{***}: p < 0.01$.